1-1-1996

Are there diagnostic alternatives to the IQ-reading discrepancy? : evaluation of assessment techniques for identifying reading disabled college students.

Cheryl A. Cisero
University of Massachusetts Amherst

ARE THERE DIAGNOSTIC ALTERNATIVES TO THE IQ-READING DISCREPANCY?: EVALUATION OF ASSESSMENT TECHNIQUES FOR IDENTIFYING READING DISABLED COLLEGE STUDENTS

A Dissertation Presented

by

CHERYL A. CISERO

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

February 1996

Psychology

ARE THERE DIAGNOSTIC ALTERNATIVES TO THE IQ-READING DISCREPANCY?: EVALUATION OF ASSESSMENT TECHNIQUES FOR IDENTIFYING READING DISABLED COLLEGE STUDENTS

A Dissertation Presented

by

CHERYL A. CISERO

Approved as to style and content by:

James M. Royer, Chair

David Arnold, Member

Ronald K. Hambleton, Member

Keith Rayner, Member

Melinda A. Novak, Chair
Psychology Department

ACKNOWLEDGEMENTS

I would like to express my gratitude to several people. First and foremost, I am indebted to my advisor and mentor, James M. Royer, not only for his wisdom, guidance, and support on this project and throughout my graduate training, but also for showing me, through his example, what it means to be a mentor, a role I could only hope to fill someday as well as he does.

I would also like to thank those involved in the project. I would like to extend my gratitude to committee members, David Arnold, Ronald Hambleton, and Keith Rayner for their valuable time and input. I am also thankful to everyone at Westfield State College who played a part in this project: Drs. Horace Marchant and Stanley Jackson for their guidance and input in the data collection process, Daniel Neuner and Brian Masciadrelli for collecting data, the professors who allowed their students to participate for extra credit, the Disabled Student Services staff for their valuable time and input regarding the learning disabled sample, and the students. Thanks are also due to Loel Tronsky for his time and effort involved in data analysis.

Finally, I would like to thank my husband, Michael Durwin, for his unending support and encouragement during this project and throughout my schooling, and my parents for their encouragement, support, and pride in my accomplishments.

ABSTRACT

ARE THERE DIAGNOSTIC ALTERNATIVES TO THE IQ-READING DISCREPANCY?: EVALUATION OF ASSESSMENT TECHNIQUES FOR IDENTIFYING READING DISABLED COLLEGE STUDENTS

FEBRUARY 1996

CHERYL A. CISERO, B.A., UNIVERSITY OF CONNECTICUT

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor James M. Royer

The current approach to identifying specific reading disability is plagued with problems. The most common diagnostic procedure, called the IQ-achievement discrepancy, involves establishing that a student's reading performance on standardized achievement tests is significantly below what would be expected from his/her IQ. This approach is unreliable with respect to diagnosis and uninformative with respect to prescriptives for remediation. An approach is needed that can provide reliable diagnosis and can indicate the deficient skills that could be targeted for remediation.

The purpose of the present research was to evaluate alternatives to the IQ-reading discrepancy for identifying reading disabled college students. Specifically, the question was whether reading disabled and nondisabled college students could be differentiated using the Computer-based Academic Assessment System (CAAS) and a measure of listening and reading comprehension called the Sentence

Verification Technique. College students recruited from Disabled Students Services and nondisabled introductory psychology students at the same college were given SVT tests and elementary-level and adult-level CAAS reading batteries. After all data was collected and prior to data analysis, students in the disabled sample were classified as having a reading disability, generalized learning disability, or other disabilities on the basis of various sources of information.

The requirements of a diagnostic technique for identifying reading disability were used as a framework for evaluating SVT and CAAS techniques. Multivariate analyses of variance were used to evaluate each of the techniques alone, and discriminant analyses were used to evaluate the techniques in combination in meeting the following requirements: 1) differentiating disabled from nondisabled students, 2) differentiating reading disabled students from nondisabled students and from students with other disabilities, 3) differentiating among disabled students with different types of problems, and 4) identifying individual patterns of performance that indicate a reading disability. Results suggested that SVT and CAAS techniques were generally able to make the above distinctions with the CAAS technique appearing to be more effective. Reasons for why SVT may have been less successful are provided in the discussion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Several laws (PL 94-142, PL 99-457, PL 101-476) have been passed within the last two decades which mandate that students with disabilities not be discriminated against on the basis of their disabilities, and more importantly, that they are entitled to individual special services. Learning disabilities are one category of disabilities protected by this legislation.

As a consequence of the legislation mentioned above, professionals in the field of education had been faced with the dilemma of how to identify learning disabilities in order to provide special services to learning disabled students. The most logical place to look for a way to identify learning disabilities was, of course, the legal definition of a learning disability. According to the Education for All Handicapped Children Act (PL 94-142),

> specific learning disability means a disorder in one or more of the basic psychological processes involved in understanding or in using language spoken or written, which may manifest itself in an imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations. The term includes such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. The term does not include children who have learning problems which are primarily the result of visual, hearing, or motor handicaps, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage (cited in Stanovich, 1991b, p. 9).

Of particular importance in this definition is the exclusion of mental retardation as a contributing factor in learning problems. The exclusion of mental retardation implies that students with learning disabilities have "normal" levels of intelligence and that their learning problem, whether reading, writing, or mathematics, is not attributable to inadequate intelligence. This notion of a learning problem despite adequate intelligence was accepted by practitioners as the key defining feature of a learning disability and was immediately transformed into an IQ-achievement discrepancy for the purpose of identifying learning disabled students. That is, a student would be identified as learning disabled if his or her achievement in one or more academic areas (as measured by standardized achievement tests) is well below what would be expected given his or her IQ.

The IQ-achievement discrepancy offered a solution to the issue of how to identify learning disabled students. The problem with this approach, however, is that IQ was accepted as a benchmark of aptitude for measuring an achievement discrepancy without critical evaluation of the research evidence (Stanovich, 1991a, 1991b). Since the adoption of the IQ-achievement discrepancy, evidence has accumulated to challenge the adequacy of the discrepancy method on theoretical, empirical, logical, statistical, and practical grounds (e.g., Evans, 1990; Fletcher et al., 1994;

Morrison & Siegel, 1991; Reynolds, 1981, 1985; Shepard, 1980; Siegel, 1989, 1992; Siegel & Heaven, 1986; Stanovich, 1991a, 1991b; Stanovich & Siegel, 1994).

The inadequacy of the IQ-achievement discrepancy has led to a call for more appropriate procedures for identifying learning disabilities. With respect to specific reading disability, researchers have emphasized the need for diagnostic procedures that are more educationally relevant than the IQ-achievement discrepancy (e.g., Siegel, 1988, 1989; Stanovich, 1991a, 1991b). This need for educationally relevant diagnostic procedures has not only been felt at the elementary and secondary levels, but at the postsecondary level as well. Recent growth in the number of reading disabled students entering postsecondary institutions (Lewin, 1995; Vogel, 1982) and a lack of useful diagnostic tools specifically normed for college students (Woods, Sedlacek, & Boyer, 1990) have prompted a search for diagnostic procedures that would be appropriate for identifying reading disabled college students.

This dissertation explores the diagnostic value of educationally relevant assessment techniques for identifying specific reading disability in college students. Educationally relevant diagnostic procedures are considered here to be those that measure the particular reading skills that have been identified by a model of reading disability to be deficient in reading disabled students. It has been suggested that disabled readers have a core deficit in phonological processing (e.g., Stanovich 1988a, 1988b, 1993) that affects their phonological awareness skills, word recognition, and even reading comprehension. Therefore, techniques that assess these reading skills may be viable alternatives to the IQ-reading achievement discrepancy for identifying specific reading disability in college students.

The dissertation begins with a brief examination of the practical problems with using the IQ-reading achievement discrepancy method for identifying students with a specific reading disabilities. The theoretical and statistical problems, although equally damaging to the IQ-achievement discrepancy method, will not be discussed. The emphasis here is on practical shortcomings of the IQ-achievement discrepancy procedure for identifying specific reading disabilities. The next section of the dissertation describes a model advanced by Stanovich (1988a, 1988b, 1993) called the phonological-core variable-difference model that provides a way to conceptualize specific reading disability. In the section that follows, aspects of the reading task which are consistent with this model are discussed as potential areas on which to develop diagnostic procedures. The subsequent section examines evidence that techniques based on reading processes targeted by the phonological-core model would be able to distinguish reading disabled and nondisabled college students. The final section includes an overview of the present research.

Practical Problems with the IQ-Reading Discrepancy

The IQ-reading achievement discrepancy approach has several practical problems that limit its diagnostic and prescriptive value. One problem is the tremendous variability among practitioners in implementing the discrepancy method. For instance, a survey of diagnostic practices among State Departments of Education has indicated enormous variation with respect to the IQ cutoff for establishing "normal" intelligence, the degree of achievement discrepancy, and the methods of calculating a discrepancy between IQ and achievement (Frankenberger & Fronzaglio, 1991).

The problem of variability in criteria and procedures is not limited to elementary and secondary education, but is also found at the postsecondary level. Woods et al. (1990) recently reported the results of a survey of 13 large public universities that provide services to learning disabled students. They found that only about half of the universities used standardized diagnostic information. Moreover, little consistency among the 13 universities was found in the types of standardized intelligence and achievement tests used, and information from the standardized tests was often supplemented by idiosyncratic data sources such as grade point averages, letters, writing samples, and interviews.

A consequence of all the variability in criteria and procedures for implementing the discrepancy method is the possibility of inconsistent diagnoses. Given one set of IQ and achievement tests and one set of criteria for "normal" IQ, degree of discrepancy, and method of calculating a discrepancy, a student may be diagnosed as reading disabled. However, with an entirely different set of tests and criteria, he or she may be considered nondisabled. In fact, it has been demonstrated that the number of students identified as learning disabled varies with different discrepancy methods (Epps, Ysseldyke, & Algozzine, 1983; Forness, Sinclair, & Guthrie, 1983; Lindgren, DeRenzi, & Richman, 1985).

A second problem with the IQ-reading achievement discrepancy is that the use of a discrepancy between IQ and achievement may lead to misdiagnosis. Consider the case of a student who has an above-average IQ score but is performing at an average level on a standardized reading achievement test. This student would be identified by a discrepancy method as reading disabled since his or her achievement score are discrepant from his or her IQ score (Rutter & Yule, 1975). Discrepancies between IQ and achievement, however, are to be expected statistically. That is, students with high IQ scores will tend to have achievement scores that are not as high. Therefore, the discrepancy approach may identify students as reading disabled who may have no disability at all.

One illustration of the problem of misdiagnosis when using discrepancies between IQ and achievement to identify learning disabilities is reported in a study by Gajar, Salvia, Gajria, and Salvia (1989). Thirty-three learning disabled college students were matched on average IQ to nondisabled students selected from a pool of freshmen at the same college. The two groups were then compared on the reading, math, written language, and knowledge subtests of the Woodcock-Johnson Psychoeducational Battery. Discrepancies were calculated between achievement scores predicted from IQ and actual achievement scores. The learning disabled group did indeed have significantly greater achievement discrepancies in reading, math, and written language than the nondisabled group. However, a better test of the diagnostic utility of achievement discrepancies is whether they can distinguish disabled from nondisabled students. The result of discriminant analyses indicated that overall nondisabled students were correctly classified 59% of the time and learning disabled students 41% of the time. In other words, a large proportion of nondisabled students was <u>incorrectly</u> classified as disabled according to the discrepancy method.

A final problem with the use of an IQ-reading achievement discrepancy is that standardized achievement tests provide little prescriptive information. Federal laws regarding education for disabled individuals dictate the development of effective educational plans that are tailored to meet the specific needs of the disabled student. Therefore, a diagnostic procedure not only needs to detect the existence of a reading disability but also should be able to identify which reading skills are deficient so that appropriate instruction can take place. Standardized reading achievement tests that are used to determine a discrepancy between IQ and reading cannot adequately target areas of reading difficulty. For instance, scores on reading achievement tests reveal that a reading-disabled student performs poorly on word recognition or comprehension subtests compared to the norm, but are unable to indicate the specific nature of the reading problem (Royer & Sinatra, 1994). Therefore, the information provided by standardized tests is often of little use in planning individualized instruction for reading disabled students.

In summary, the problems mentioned above suggest that the IQ-reading achievement discrepancy has poor diagnostic properties and little prescriptive value. First, the wide variation among practitioners in the implementation of the IQ-reading discrepancy may lead to inconsistent diagnoses. That is, it is very likely that a student would be identified as reading disabled by one set of criteria and nondisabled by another. An additional diagnostic problem concerns the notion of discrepancy. The discrepancy method may identify students as disabled who are actually nondisabled simply because their "normal" achievement scores are discrepant from their above-average IQs. Finally, standardized test scores cannot provide information that would be valuable for the placement or treatment of reading disabled students.

The limitations of the discrepancy approach discussed in this section indicate that a new approach to identifying a specific reading disability is needed. Specifically, there is a need for techniques that approach the problem of diagnosing reading disability from the perspective of identifying and isolating the particular deficits that are characteristic of reading disabled students. In other words, we need to develop techniques based on a framework of knowledge about the reading performance of disabled readers. The phonological-core variable-difference model of Stanovich (1988a, 1988b, 1993), to be discussed in the next section, provides such a framework.

The Phonological-Core Variable Difference Model of Reading Disability

Stanovich's (1988a, 1988b, 1993) phonological-core variable-difference model provides a framework for conceptualizing the cognitive deficits of disabled readers and how they differ from the profile of deficits exhibited by "garden-variety" (Gough & Tunmer, 1986) poor readers (i.e. those whose reading performance is consistent with their IQ). Before discussing the details of the phonological-core variable-difference model, it is important to explain the basis for this model.

The model rests on the assumption of specificity which is inherent in almost all definitions of specific reading disability. The assumption of specificity states that a person with a specific reading disability "has a brain/cognitive deficit that is reasonably specific to the reading task" (Stanovich, 1988b, p. 155). That is, the deficits of reading disabled students should not extend into other domains of cognitive functioning. If disabled readers exhibited a wide range of cognitive deficits, these global deficits would lower their performance on intelligence tests, thereby reducing the IQ-achievement discrepancy. As a consequence, the students would no longer be considered reading disabled by the IQ-achievement discrepancy definition but garden-variety poor readers whose reading performance is consistent with their IQ (Stanovich, 1988b).

Based on the assumption of specificity, the key deficit in specific reading disability must be a modular process (see Fodor, 1983; Stanovich, 1988b, 1990). In other words, it must be a domain-specific process that is not controlled by higher-level processes or strongly interactive with them

(Stanovich, 1988b). The reason is that modular processes can fail without necessarily hindering higher-level processes that are needed for performance on intelligence tests (Stanovich, 1988b). If higher-level processes, such as language comprehension, metacognition, and strategic functioning, were the locus for the deficit in specific reading disability, then the definition of a disabled reader having low reading performance that is discrepant from his or her IQ would no longer hold.

Consistent with the requirement of the assumption of specificity that the key deficit in specific reading disability must be modular, many researchers have isolated the locus of reading disability at the word recognition level (e.g., Adams & Bruck, 1993; Bruck, 1990; Bruck & Treiman, 1990; Compton & Carlisle, 1994; Gough & Tunmer, 1986; Morrison, 1984, 1987; Siegel, 1985, 1988; Siegel & Faux, 1989; Stanovich, 1986). Moreover, a growing body of evidence has identified various aspects of phonological processing as responsible for the impaired word recognition module of disabled readers. One aspect of phonological processing that is critical for skilled word recognition is decoding, or the application of grapheme-phoneme correspondence rules. Decoding ability is typically measured by tasks that involve pseudowords since they require the subject to use grapheme-phoneme rules in order to decode the printed nonword into sound. Disabled readers have exhibited deficits in decoding on tasks such as pseudoword naming, pseudoword spelling, and deciding which of two printed pseudowords sounds like a word (e.g., Baddeley, Ellis, Miles, & Lewis, 1982; Ben-Dror, Pollatsek, & Scarpati, 1991; Bruck, 1988; DiBenedetto, Richardson, & Kochnower, 1983; Kochnower, Richardson, & DiBenedetto, 1983; Manis, Szeszulski, Holt, & Graves, 1988, 1990; Manis, Szeszulski, Howell, & Horn, 1986; Olson, Kliegl, Davidson, & Foltz, 1985; Snowling, 1980, 1981).

An equally important phonological processing skill for word recognition is phonological awareness, or the knowledge that spoken language can be decomposed into smaller units of sound. There is substantial evidence that phonological awareness is critical for the acquisition of grapheme-phoneme correspondences and subsequent word recognition skill (Bradley & Bryant, 1985; Bryant, Maclean, & Bradley, 1990; Bryant, Maclean, Bradley, & Crossland, 1990; Lundberg, Frost, & Peterson, 1988; Maclean, Bryant, & Bradley, 1987; Perfetti, Beck, Bell, & Hughes, 1987). Therefore, the evidence that disabled readers are deficient in phonological awareness (e.g., Bruck, 1992; Bruck & Treiman, 1990; Ellis & Large, 1987; Manis et al., 1988; Pennington et al., 1990) is not surprising.

Based on the empirical evidence and on the assumption of specificity, Stanovich's (1988a, 1988b, 1993) phonological-core variable-difference model highlights a core deficit in phonological processing as the basis of specific reading disability. The phonological problems of disabled readers are, according to the model, shared by garden-variety poor readers. The difference between these two types of readers is that disabled readers have cognitive deficits that are relatively specific to phonological processing, while poor readers exhibit a wide variety of cognitive deficits that include, but are not limited to, phonological processing problems. Therefore, the term "variable-difference" in the model's name refers to the fact that the actual cognitive deficits displayed by students with reading problems will be variable depending on whether the student is reading disabled or a poor reader. On one end of a continuum of reading disability, deficits will be located in the phonological core (characterizing specific reading disability), and the deficits will increase in number along the continuum toward the garden-variety poor reader at the other end of the continuum who will have a multitude of cognitive deficits (Stanovich, 1993).

It is evident that Stanovich's model of reading disability emphasizes phonological processing as the root of specific reading disability. One limitation of this model, however, is that it ignores the existence of a small number of reading disabled students with visual/perceptual deficits or orthographic processing deficits (problems dealing with irregularities in orthography). In fact, Stanovich (1988a) admits that an argument for phonological deficits as the sole basis of a reading disability is an oversimplification.

Stanovich's (1993) model of reading disability does not include visual/perceptual and orthographic deficits because the research evidence is still unclear on the role of visual/perceptual and orthographic deficits in specific reading disability. For instance, there is no consistency among studies concerning the type of visual/perceptual problem that is characteristic of reading disability. The types of deficits that have been found in reading disabled subjects have ranged from deficits in the transient visual system (a flicker or motion system that transmits information about a change in the visual stimulus) (Lovegrove, Martin, & Slaghuis, 1986; Slaghuis & Lovegrove, 1985), oculomotor deficits (Stein & Fowler, 1982), scotopic sensitivity syndrome (Irlen, 1991), and a problem attending to letters in foveal vision (Rayner, Murphy, Henderson, & Pollatsek, 1989). Moreover, researchers do not seem to agree on whether visual/perceptual problems are a cause of reading disability (e.g., Lovegrove et al., 1986; Slaghuis & Lovegrove, 1985; Hulme, 1988). One reason for the difficulty in determining whether visual/perceptual processes are a cause of reading disability is that deficits in visual/perceptual processes often co-occur with phonological deficits (Stanovich, 1992).

Like visual/perceptual deficits, orthographic processing deficits have been difficult to isolate from phonological deficits. Studies that have examined phonological and orthographic processes in reading disabled individuals have found that only a very small number of reading disabled individuals has purely orthographic deficits (Manis, Szeszulski, Holt, & Graves, 1988, 1990).

Given the evidence that phonological processing deficits occur in a majority of reading disabled individuals and given the present lack of definitive evidence for visual/perceptual deficits or orthographic deficits that exist apart from phonological processing problems, it is fitting that the current model of reading disability emphasizes phonological processing deficits as the basis of specific reading disability. It seems logical, then, that the development of diagnostic techniques for identifying specific reading disability should also, at present, focus on phonological processes. The next section discusses the phonological processes underlying reading that could be used as the basis for diagnostic techniques.

Implications of the Phonological-Core Variable-Difference Model for the Development of Diagnostic Techniques

The phonological-core variable-difference model discussed in the previous section not only provides a way to understand the mass of research evidence regarding the cognitive deficits of disabled readers (as well as those of poor readers), but also provides a suitable framework from which to develop techniques that could be used to diagnose disabled readers. If disabled readers indeed have a core of phonological deficits, as the model and the research evidence suggest, then aspects of the reading task that involve or are influenced by phonological processing can be identified as sources of cognitive breakdown. These sources of difficulty would logically be the areas on which to develop diagnostic procedures.

Sources of Difficulty in Reading

One source of difficulty relevant to reading that involves phonological processing is phonological awareness. Phonological awareness is a knowledge of the component sounds of one's language that enables children to learn to read. This knowledge allows a beginning reader to discover the alphabetic principle that printed letters are represented by sounds (Liberman & Liberman, 1990). Children who discover the alphabetic principle are able to learn to read unfamiliar words by applying their knowledge of grapheme-phoneme correspondences, or decoding.

In contrast, children who do not acquire a sufficient level of phonological awareness may be at risk for developing reading problems. A lack of phonological sensitivity may hinder the acquisition of grapheme-phoneme correspondences and subsequent word decoding skill. Beginning readers who laboriously decode words are unable to acquire efficient word recognition skill. Word recognition/decoding, therefore, is another source of difficulty in reading.

Because word recognition remains effortful for children who lack decoding skill, reading is a frustrating experience, which causes children to read less, thereby inhibiting further growth in their reading skills (Stanovich, 1986). As a consequence, higher-level reading processes may become another source of difficulty. For instance, inefficient word recognition processes may create a "bottleneck" in reading comprehension (Perfetti, 1985; Perfetti & Lesgold, 1977). In skilled reading, word recognition is generally very fast and nearly load-free (Perfetti, 1992; Stanovich, 1990). Fast word recognition allows a reader to hold more words in working memory (which has limited capacity and a very short duration) so that a meaningful chunk can be processed before words that were initially processed decay from working memory. Load-free word recognition requires fewer cognitive resources so that cognitive capacity can be allocated to higher-level comprehension activities. Therefore, a failure to develop fast and load-free word recognition may adversely affect reading comprehension in two ways: 1) cognitive capacity would be used for word recognition rather than comprehension processes, and 2) rapid decay of words from working memory would make it less likely that readers will be able to process strings of words into meaningful segments.

Cognitive Deficits of Reading Disabled Individuals

Phonological awareness, decoding, and reading comprehension are potential sources of difficulty in reading. These areas of difficulty, though, are not limited to disabled readers but apply to garden-variety poor readers as well. The reason is that these areas of difficulty either directly involve phonological processing or are influenced by phonological processing, which is considered a core deficit for both disabled readers and garden-variety poor readers. Phonological awareness and decoding directly involve phonological processing, and reading comprehension difficulties would result from inefficient decoding. Therefore, the cognitive deficits with respect to the reading task may be similar for disabled readers and garden-variety poor readers.

The key difference between disabled readers and garden-variety poor readers, according to the phonological-core model, is that poor readers have deficits in higher-level skills outside the domain of reading in addition to their phonological deficits. Therefore, a logical corollary of the phonological-core model (and the assumption of specificity upon which it is based) is that poor readers will have poor listening comprehension, which is a higher-level function, in addition to their poor reading comprehension. In contrast, disabled readers will have "normal" listening comprehension ability and suppressed reading comprehension. That is, the reading comprehension ability of disabled readers will be discrepant from their "normal" listening comprehension ability.

To sum up, according to the phonological-core model, disabled readers would exhibit cognitive deficits in phonological awareness, decoding, and reading comprehension, but not listening comprehension. Therefore, it is possible that techniques designed to measure these skills may be able to identify disabled readers. The next section examines evidence which indicates that phonological awareness, decoding, and listening and reading comprehension can distinguish reading disabled and nondisabled adults.

Support for the Use of Theory-based Techniques in Identifying Specific Reading Disability in Adults

In the previous section, it was suggested that a discrepancy between listening comprehension and reading comprehension and deficits in phonological awareness and decoding would be characteristic of disabled readers, and would thus be the areas upon which diagnostic techniques should be based. The possibility that skills such as phonological awareness, decoding, and listening and reading comprehension could be used to diagnose a specific reading disability is not a new one. Stanovich (1991a, 1991b) has argued that a discrepancy between listening comprehension and reading comprehension would have been a more appropriate method of identifying disabled readers than the IQ-reading discrepancy. Likewise, Siegel (1988, 1989) has suggested that a decoding deficit should be the key defining feature of specific reading disability.

Empirical support for using techniques that measure particular reading skills to diagnose specific reading disability has mainly been obtained in studies with children. For instance, children with a specific reading disability have been found to have listening comprehension comparable to nondisabled students but significantly poorer reading comprehension (Aaron, 1991; Aaron, Kuchta, & Grapenthin, 1988; Spring & French, 1990). The performance of reading disabled children on a variety of phonological awareness tasks has also been found to be poor relative to younger nondisabled readers who are matched to disabled readers on reading level (i.e. reading-age matched controls) (Bruck & Treiman, 1990; Ellis & Large, 1987; Manis et al., 1988). Moreover, there is substantial evidence that reading disabled children have poorer decoding skills than reading-age matched controls. For instance, disabled readers perform significantly worse than younger reading-age matched students on pseudoword reading (Kochnower et al., 1983; Manis et al., 1988, 1990; Snowling, 1981) and pseudoword spelling (Manis et al., 1988). They also perform more poorly on other tasks that require decoding, such as a pseudoword verification task in which subjects decide whether a pseudoword that was pronounced is the same as one presented by computer (Manis et al., 1988, 1990), a pseudoword matching task in which subjects decide whether two similarly spelled pseudowords have the same sound (Manis et al., 1988, 1990), and a phonetic task that requires subjects to determine which of two pseudowords sounds like a word (Olson, 1985).

In contrast, little is known about the phonological awareness, decoding, and comprehension skills of adults. It is possible that these skills, while deficient in reading disabled children, no longer present a problem for older disabled readers. Therefore, diagnostic techniques designed to assess these skills would not be useful. However, the limited amount of evidence appears to indicate that phonological awareness, decoding, and listening and reading comprehension measures are capable of distinguishing reading disabled and nondisabled adults.

Phonological Awareness Evidence

To date, there are two studies that have examined phonological awareness capabilities in reading disabled adults. The results of both studies suggest that reading disabled adults have poorer phonological awareness skills relative to their peers and to younger nondisabled readers.

Pennington, Van Orden, Smith, Green, and Haith (1990) compared the phonological awareness performance of adult disabled readers with chronological-age matched and reading-age matched control groups. Reading disabled adults were recruited from a reading disability program at a local college or from families with three-generation histories of reading disability. Disabled readers were matched to chronological-age controls on age and gender, and to younger reading-age controls on performance on the Reading Recognition subtest of the Peabody Individual Achievement Test.

Two types of "pig latin" tasks were used to assess phonological awareness. Pig latin production required subjects to produce a pig latin form from a real target word (e.g., bank: "ank-bay"), and pig latin recognition required subjects to recognize the correct pig latin form for a target word. Accuracy and response time were recorded. Reading disabled adults were significantly less accurate and slower than both chronological-age matched and reading-age matched control groups.

Similar to the study by Pennington et al. (1990), Bruck (1992) compared the phonological awareness performance of reading disabled adults to both same age peers and younger nondisabled readers. Reading disabled subjects were adults between ages 19 and 27 who were diagnosed with specific reading disability as children. Their performance on phonological awareness tasks was compared to that of nondisabled college students and to that of younger nondisabled readers in grades 1, 2, and 3. The phonological awareness tasks used were syllable and phoneme counting of pseudowords and deletion of initial phoneme or final phoneme in pseudowords. Reading disabled adults made significantly more errors in phoneme counting and deletion of final phonemes than both college students and younger nondisabled readers.

In summary, the available research evidence indicates that the deficit in phonological awareness persists into adulthood. Therefore, tasks measuring phonological awareness may be suitable for identifying disabled readers at the adult level.

Word Recognition/Decoding Evidence

The results of a few studies that have been conducted with college students suggest that decoding problems are a key feature of reading disability even at the college level. Evidence indicates that reading disabled college students have poor decoding skills relative to their nondisabled peers and to younger nondisabled readers. There is also evidence that disabled readers have deficits in higher-level components of reading as well as in word recognition.

Assessment of the Word Recognition Component. Ben-Dror et al. (1991) examined the word recognition skills of reading disabled college students relative to chronological-age matched controls and reading-age matched controls. Reading disabled students had been previously diagnosed as having a specific reading disability. The reading disabled group had a mean full-scale IQ of 106 on the Wechsler Adult Intelligence Scale and scored below the 40th percentile on the Woodcock Reading Mastery Test. Chronological-age (CA) matched controls were matched to reading disabled students on age, gender, and IQ. Reading-age (RA) matched controls were matched to disabled readers on word identification level as measured by the Woodcock Reading Mastery Test. Subjects were given words and pseudowords to read aloud. Reading disabled subjects were significantly slower and less accurate at naming both words and pseudowords than CA and RA controls.

The results of the study by Ben-Dror et al. (1991) are consistent with the phonological-core model and with research on reading disabled children. That is, disabled readers, even as adults, exhibit great difficulty in phonological processing at the word recognition level as measured by decoding of words and pseudowords.

Assessment of Word Recognition and Other Component Skills. An implication of the phonological core model that was discussed in the previous section is that deficits in the word recognition component of reading may lead to problems in higher-level reading components. The results of two studies by Cisero and colleagues suggest that reading disabled college students have problems in higher-level components of reading as well as severe deficits in word recognition (Cisero, Royer, Marchant, & Jackson, 1995; Cisero, Royer, Marchant, & Wint, 1994).

Cisero et al. (1994, 1995) have examined the component reading skills of reading disabled college students using the Computer-based Academic Assessment System (CAAS). The CAAS system, developed by Royer, is a computer-based system that measures the speed and accuracy of performance on tasks designed to assess component processes involved in reading (Royer & Sinatra, 1994; Sinatra & Royer, 1993). In all tasks, stimuli are presented by computer and subjects make a response into a microphone. The computer records the latency of the response and the examiner records the accuracy of the response by pressing a correct or incorrect button on a box that is connected to the computer.

Tasks included in the CAAS system are: a simple response time task, a letter naming task, word and pseudoword naming tasks, a category match task, and two sentence tasks that measure syntactic and semantic processing. The simple response time task is a non-reading task designed to measure a subject's response time to non-verbal stimuli. Subjects are presented with "***" or "+++" and respond by saying "star" or "plus." In the letter, word, and pseudoword naming tasks, the stimulus appears on the computer screen and the subject must say the letter or word, or pronounce the pseudoword. In the category match task, subjects are informed of the categories that will be used, and are then presented with pairs of words, and must decide whether or not they belong to the same category. The syntax and semantics tasks present subjects with sentences that have a blank in them and a word appearing above and below the blank. Subjects need to decide which of the two words best fits the sentence. The two words vary in syntactic appropriateness (e.g., The boy eat/ate his dinner.) in the syntax task and semantic appropriateness (e.g., The boy ate/drank his milk.) in the semantics task. Stimuli in the word, category, syntax, and semantics tasks were considered to be at a fourth grade difficulty level.

The study by Cisero, Royer, Marchant, and Wint (1994) examined the CAAS performance of nondisabled students and students who were identified as reading disabled or undifferentiated learning disabled in order to determine whether different patterns of performance would emerge as a function of type of disability. Twenty-eight students from a college in western Massachusetts were recruited from Disabled Student Services at the college. Nineteen were identified as having a specific reading disability and 9 were identified as having an undifferentiated learning disability. Forty students from an introductory psychology class at the same college served as a nondisabled comparison group.

Subjects were individually administered the CAAS tasks described above, except that this version of the CAAS system used a variation of the Posner letter match task (e.g., Posner, Boies, Eichelman, & Taylor, 1969) rather than letter naming. In the letter match task, subjects are given a pair of letters and must decide if they have the same name (e.g., same name: Aa, AA; different: Ab, AB). Accuracy and response time data were collected for all tasks.

Nondisabled students were most accurate on all tasks, followed by reading disabled students, and then learning disabled students. Nondisabled students were also significantly faster overall than the two disabled groups. The most interesting result, though, was that the two disabled groups showed significantly different patterns of response time performance on the tasks.

For ease of interpreting the differential response time performance of reading disabled and learning disabled students on CAAS tasks, response time scores were converted into effect sizes (which provide a Z score indication of where the average disabled student would score if he or she were in the nondisabled group). The effect sizes were then transformed into percentile scores so that nondisabled students were defined as being at the 50th percentile. Figure 1 displays the percentile scores of reading disabled and learning disabled students on all CAAS tasks relative to nondisabled students.

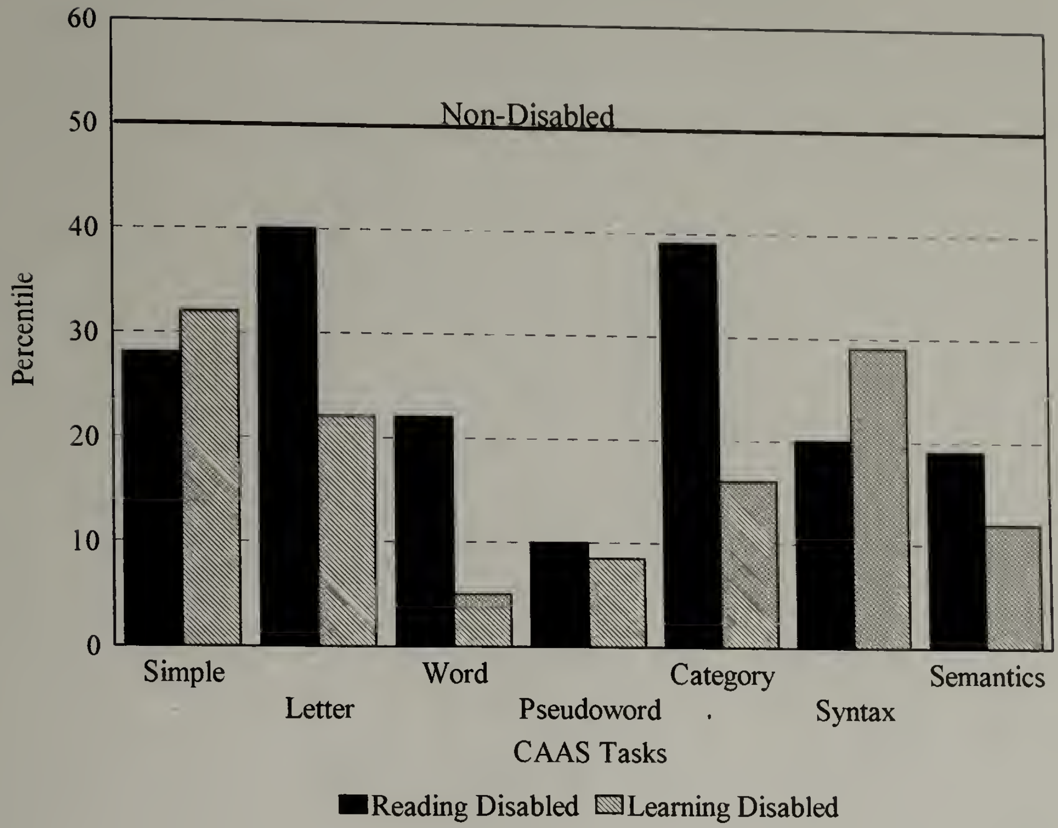

Figure 1. Performance of reading disabled and learning disabled college students on CAAS tasks relative to nondisabled college students. Nondisabled students are represented by the solid line at the 50th percentile.

As can be seen in the figure, learning disabled students were generally slow across all tasks, while reading disabled students performed comparably to nondisabled peers on some tasks but substantially slower on other tasks. Specifically, learning disabled subjects scored at or below the 30th percentile (relative to nondisabled subjects) on every task, with word and pseudoword naming performance being the lowest. In contrast, reading disabled subjects scored at about the 40th percentile on the letter match and category match tasks, indicating performance that is similar to nondisabled students. However, the performance of reading disabled subjects on word naming, syntax, and semantics tasks was at approximately the 20th percentile, and pseudoword naming was at the 10th percentile.

The differential pattern of performance for the two disabled groups suggests a more specialized area of cognitive deficits for the reading disabled subjects than for the general learning disabled subjects. Consistent with the phonological-core model, reading disabled and general learning disabled subjects both have deficits in the phonological core as measured by word and pseudoword naming. These phonological deficits in the area of word identification may also be responsible for the poor performance of reading disabled subjects on the syntax and semantics sentence tasks. While this may also be said for the general learning disabled students, it is important to keep in mind that the learning disabled subjects were slow relative to nondisabled students on *all* tasks, including the simple response time task which has nothing to do with reading. The learning disabled pattern of performance, then, indicates a general "cognitive sluggishness" (Royer & Sinatra, 1994). The different patterns of performance for nondisabled, reading disabled, and general learning disabled students provides support for the ability of the CAAS system to identify disabled readers.

A second study by Cisero, Royer, Marchant, and Jackson (1995) further examined the cognitive profiles of reading disabled college students on the CAAS system. Subjects were 8 students identified as reading-disabled by Disabled Student Services at a college in western Massachusetts. Thirty-five nondisabled students from an introductory psychology course at the same college served as a comparison group. Subjects were administered the same CAAS tasks as in the previous study mentioned above, with two exceptions. A letter naming task was substituted for the Posner letter match task and the syntax task was not included. Subjects were also given the adult version of the CAAS battery which consisted of word, pseudoword, category, and semantics tasks analogous to the original elementary level tasks (which contained stimuli suitable for fourth graders) except with more difficult stimuli.

A transformation procedure was used to combine accuracy and response time into a single index (Sinatra & Royer, 1995). Effect sizes were calculated for the combined index, and were then converted into percentile scores where nondisabled performance was defined as being at the 50th percentile. Subjects were sorted into one of three groups, each of which represented a different profile of performance relative to the nondisabled group.

The authors report the profiles of performance of three reading disabled students who are representative of each category of performance. The performance of subject SH, shown in Figure 2, represents a compensatory profile. That is, the student has learned to compensate to some extent for her disability. This is indicated by performance on the simple, letter, and elementary level word and pseudoword tasks that is comparable to nondisabled peers. The fact that she performs well on relatively simple material may be a consequence of receiving remedial instruction for her disability since the 6th grade. Her difficulty surfaces on the more complex elementary level category and semantics tasks, where word identification is only part of the task requirements. With respect to all adult-level tasks, where vocabulary is more difficult and perhaps less practiced, SH performs below the 5th percentile. It should be noted that SH's performance on CAAS tasks is consistent with reports of

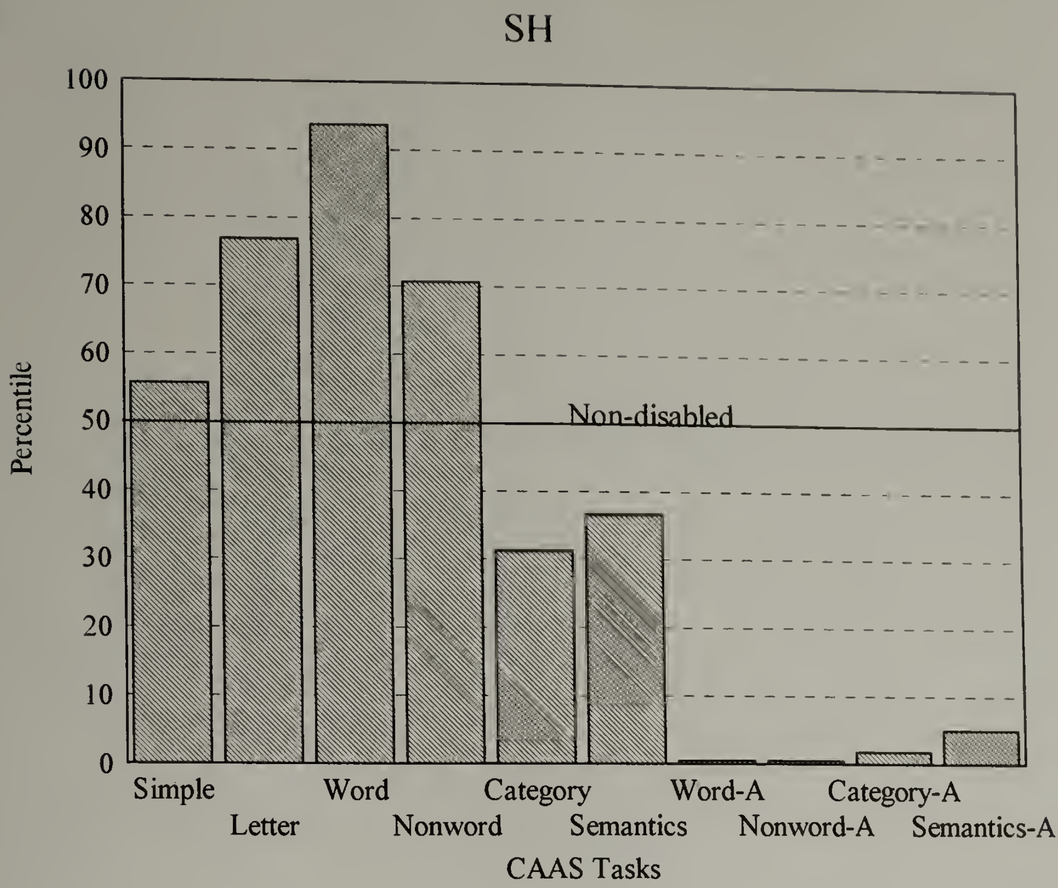

Figure 2. Performance of reading disabled college student, SH, on CAAS tasks relative to nondisabled college students. Nondisabled students are represented by the solid line at the 50th percentile.

her below average reading achievement for her age on the Woodcock-Johnson Psycho-Educational Battery.

The performance of subject EG in Figure 3 represents a severe reading disability profile. She performs at the level of nondisabled peers on the simple and letter tasks. However, her performance on all other elementary level and adult-level reading tasks is at or below the 10th percentile. This is consistent with findings from standardized tests that indicate below average word reading, spelling, and reading rate. Like SH, EG has received remediation for her reading disability since elementary school. However, her CAAS profile indicates that her reading skills, even with respect to familiar vocabulary, are still severely impaired.

The final profile displayed in Figure 4 is that of CM. CM's profile appears to be relatively "normal." That is, her performance is comparable to nondisabled peers on almost all elementary level and adult-level tasks. In particular, her performance on the letter and word tasks is consistent with average to above-average performance on the Letter-Word Identification and Word Attack (words in isolaticn) subtests of the Woodcock-Johnson Psycho-Educational Battery.

CM's "normal" CAAS profile is inconsistent with her reading-disabled diagnosis. CM's diagnosis was based on a discrepancy between her IQ and achievement test scores. Her IQ score (139) is over two standard deviations above the

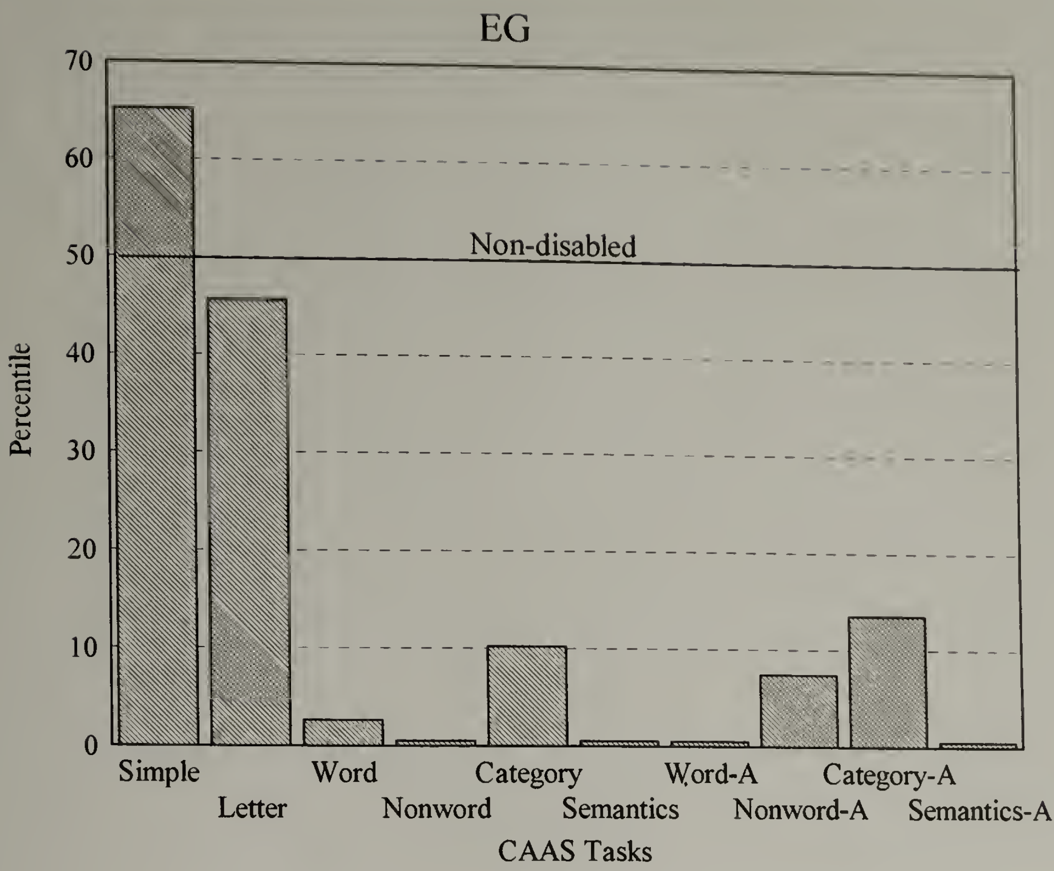

Figure 3. Performance of reading disabled college student, EG, on CAAS tasks relative to nondisabled college students. Nondisabled students are represented by the solid line at the 50th percentile.

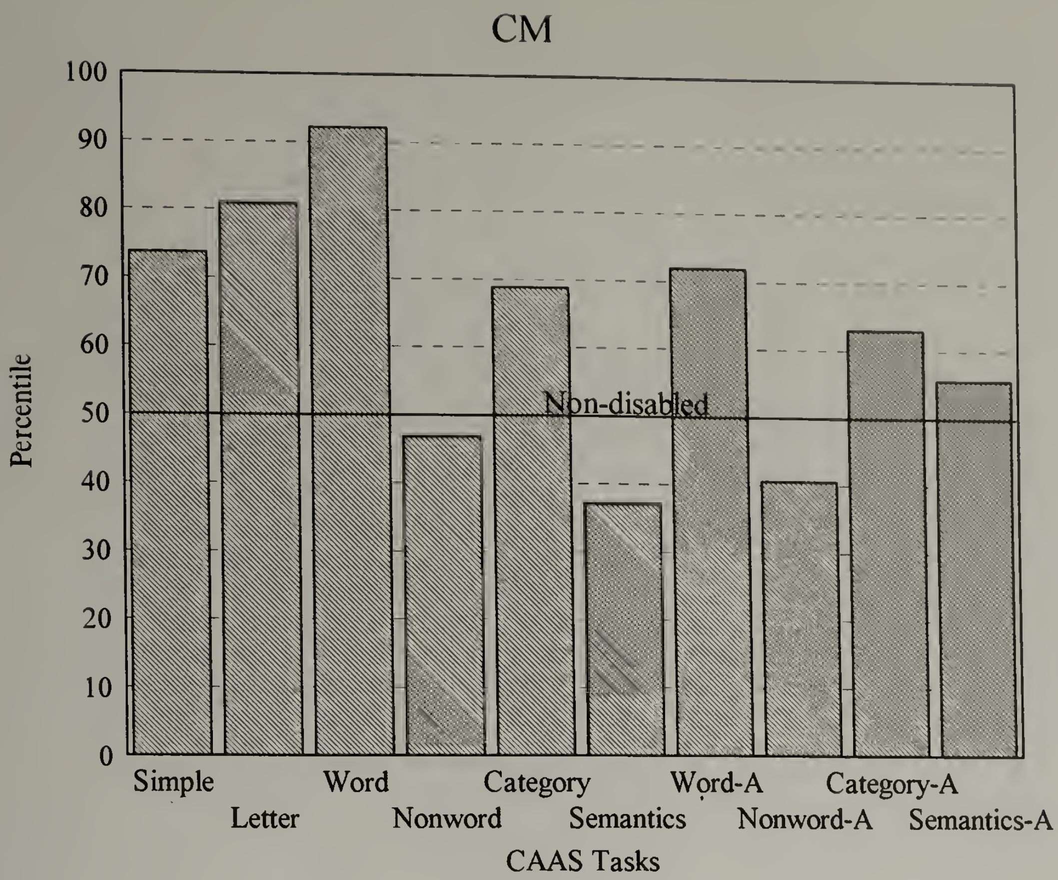

Figure 4. Performance of reading disabled college student, CM, on CAAS tasks relative to nondisabled college students. Nondisabled students are represented by the solid line at the 50th percentile.

mean, while her achievement scores in listening, reading comprehension, writing, and applied mathematics are average (as the CAAS profile is). It is expected that students with high IQs will tend to have achievement scores that are not as high. Therefore, CM's average achievement scores, although discrepant from her superior IQ, are not surprising. Moreover, the fact that her achievement level is average suggests that she may not be reading-disabled at all. The case of CM supports the argument made earlier in the dissertation that the IQ-reading achievement discrepancy may misdiagnose students with high IQ scores and average reading skills as having a specific reading ability.

The fact that the CAAS profiles of reading disabled students were consistent with official documentation of their reading performance indicates that the CAAS system may be useful for identifying college-age disabled readers. Moreover, the different profiles of performance exhibited by reading disabled students suggests potential usefulness of CAAS results for providing information about individualized intervention, a feature that the IQ-reading discrepancy does not have.

In sum, research indicates that reading disabled college students have a deficit in decoding as measured by word and pseudoword naming tasks. Moreover, studies examining the diagnostic potential of the CAAS system suggest that assessing a number of component reading skills, rather than just word and pseudoword decoding, may provide more information about the areas of deficits that disabled readers exhibit.

A Discrepancy between Listening Comprehension and Reading Comprehension

According to the implications of the phonological-core model, disabled readers would experience problems in reading comprehension as a result of inefficient word recognition, but would have intact listening comprehension. Therefore, one technique that could be used to identify adult disabled readers would be the measurement of a discrepancy in reading comprehension from "normal" listening comprehension (i.e. a listening comprehension-reading comprehension discrepancy). Research on the listening and reading comprehension abilities of reading disabled adults is virtually nonexistent. The only available evidence to suggest that a listening-reading comprehension discrepancy would be useful for identifying adult disabled readers comes from a study by Aaron (1987).

Aaron (1987) reports evidence that reading disabled college students have "normal" listening comprehension while poor readers do not. Reading disabled subjects were 14 college students in remedial reading classes who were reading 3 or more years below what was expected for grade level. Disabled readers were divided into two groups based on their Full-Scale IQ scores on the Wechsler Adult

Intelligence Scales. Those with a Full-Scale IQ of 95 or above were classified as having a specific reading disability since their low reading achievement was discrepant from their IQ scores. Reading disabled students with a Full-Scale IQ of 85 or below were classified as having a non-specific reading disability (i.e. garden-variety poor reader) since their poor reading achievement was consistent with their lower IQ levels. Seven nondisabled readers from undergraduate classes served as controls.

Listening comprehension was assessed as part of a battery of tests. Subjects were given the listening comprehension subtest of the Durrell Analysis of Reading Difficulty, which requires examinees to listen to passages and answer questions involving recall of details from the passage. Specific reading disabled subjects performed similarly to nondisabled controls on the listening comprehension test, while non-specific reading disabled subjects had significantly lower listening comprehension scores. This pattern is consistent with the prediction of the phonological-core model that poor readers will have global deficits while disabled readers will have deficits that are specific to reading.

A limitation of this study is that subjects were not given a reading comprehension test that was analogous to the listening comprehension test, so that direct comparisons of disabled readers' listening comprehension and reading comprehension cannot be made. However, if it can be assumed that adult disabled readers will have depressed reading comprehension scores, then the finding that reading disabled college students have "normal" listening comprehension would suggest the possibility that a listening-reading comprehension discrepancy may be useful for identifying specific reading disability at the college level.

Summary

The current state of reading disability diagnosis is in need of change. The IQ-reading achievement discrepancy, in light of much research evidence, is not an appropriate method of identifying students with a specific reading disability. Researchers have called for techniques that are able to isolate the actual cognitive deficits of disabled readers. Deficits that have been targeted by a phonological-core model of reading disability as characteristic of disabled readers are: phonological awareness, decoding ability, and reading comprehension. A fair amount of research with reading disabled children suggests that these areas are the ones on which assessment techniques should be developed. The research on the phonological awareness, decoding, and listening and reading comprehension skills of reading disabled adults, although scarce, also suggests that techniques designed to measure these skills could be used to diagnose specific reading disability in college students. The task, therefore, is to develop techniques to measure these skills and evaluate their usefulness for identifying college-age disabled readers.

Present Research

The aim of the present research is to evaluate whether two particular assessment techniques would be of value in identifying reading disabled students at the college level. Specifically, the question is whether the Computer-based Academic Assessment System (CAAS) and the Sentence Verification Technique (SVT) for measuring listening and reading comprehension (Royer, Hastings, & Hook, 1979; Royer, Kulhavy, Lee, & Peterson, 1986) would be able to distinguish disabled readers from nondisabled readers. Students from Disabled Student Services at a college in western Massachusetts and nondisabled students from undergraduate psychology courses at the same college were administered the SVT and CAAS techniques. After all data were collected, students from the disabled sample were classified as having a reading disability, generalized learning disability, or specific disability other than reading based on several sources of information. Chapter 2 outlines the methodology of the study in greater detail.

CHAPTER 2

METHOD

Subjects

Thirty-seven learning disabled students were recruited from Disabled Student Services at a college in western Massachusetts to participate in the study. A group of 42 nondisabled students at the same college participated for extra credit in their introductory psychology courses. Both the learning disabled group and the nondisabled group had a mean age of 21. The learning disabled group was also similar to the nondisabled group in ethnicity (94% Caucasian for the disabled group and 84% for the nondisabled group) and gender (63% and 64% female for the disabled and nondisabled groups, respectively).

Subjects were classified into diagnostic categories. Subjects from the nondisabled sample made up the nondisabled diagnostic category, and subjects from the learning disabled sample were classified into one of three remaining diagnostic categories. The first category, the reading disability (RD) category, contained disabled students who exhibited problems such as reading comprehension, word recognition, or decoding, or who were characterized as slow readers. The second category, called the generalized learning disability (LD) category, consisted of disabled students who had a general learning disability rather than a specific difficulty in one academic area. These were students who exhibited deficits in multiple areas of cognitive functioning, who have been identified as slow processors, or who needed untimed tests. The last category, termed "other," contained disabled students who had difficulties in areas other than reading.

Classification of disabled subjects into the RD, LD, and "other" diagnostic categories was based on several sources of information. The primary source of information was the clinical judgment of the Counselor at Disabled Student Services who works with the students on a daily basis.[1] The Counselor was asked what she thought the student's primary difficulty was based on her responsibilities of arranging tutors and suggesting modifications of the curriculum to accommodate the student's disability. Other data were used as a supplement to the Counselor's observations: 1) student's self-report of difficulty (students were asked what they thought their primary difficulty was), 2) description of difficulty from official evaluation reports (i.e. summaries provided by psychologists regarding the nature of the disability), 3) standardized IQ and achievement test scores (a Full-Scale IQ score or verbal IQ score of at least 85 and standardized reading scores at least 2 years below grade level or below the 30th percentile were needed for a subject to be considered reading disabled), and 4) history (i.e. a history of a specific learning disability was indicated by test scores from elementary school or high school which suggested a learning disability, or by reports of earlier learning problems from personal interviews with students conducted at Disabled Student Services).

The main reason for using the Counselor's observations as the primary source of data and for using the other sources as supplements is that the Counselor was the only data source that was consistently available for all students. For instance, standardized test score information was missing (either missing IQ scores, achievement scores, or missing both) for 18 of the 37 learning disabled students (48.6%). Moreover, information regarding whether there was a history of learning difficulties was missing in 24 of the 37 cases (64.9%). In fact, the description of a student's learning difficulty from summaries of psychologists' evaluations was added as a source of information (albeit a vague source of information) for classifying disabled subjects into diagnostic categories, in part, because of the overwhelming absence of standardized test information and history information.

It is worthwhile to note that an initial attempt was made to classify disabled subjects using a combination of self-report data, the description of the difficulty from the official evaluation, test scores, and history. However, due to the abundance of missing information, reliable classifications were not possible. Therefore, the author resorted to using clinical judgment as the primary information source and using information from other sources as supplements. It is acknowledged that a more ideal approach to classifying students into diagnostic categories would be one that was less subjective than the method that was used. However, the fact that classification needed to be based mainly upon clinical judgment serves to reinforce the argument made earlier in the dissertation that current diagnostic practices lack any standardized methods of identifying learning disabilities.

The author and a graduate student who was naive to the purpose of the study independently classified subjects from the learning disabled sample into diagnostic categories using information from the various sources. Disabled subjects were classified as belonging to the RD, LD, or "other" diagnostic categories if information from all available sources suggested difficulties characteristic of a particular category. Whenever information from several sources was inconsistent or when information was ambiguous, a subject was not classified.

The overall agreement in classifications between the two raters was 90.9% (where agreement by chance would be 34.7%). Given the reasonably high agreement in classifications between the raters, only the classifications of the author were used in analyses. Seven disabled subjects were classified as reading disabled (RD), 10 as generalized learning disabled (LD), 17 as other, and 3 subjects could not be classified. Of the students in the "other" category, 5 were identified as having a math disability, 3 had Attention Deficit Disorder, 5 had general knowledge problems or difficulties grasping meanings and ideas, 1 had a long-term memory problem, 1 had a visual perception problem, and 2 had uncertain diagnoses.

Materials

Listening and Reading Comprehension Tests

Listening and reading comprehension was measured using a technique developed by Royer and colleagues (Royer, Hastings, & Hook, 1979; Royer, Kulhavy, Lee, & Peterson, 1986) called the Sentence Verification Technique (SVT). An examinee listens to or reads a passage, and then listens to or reads each test sentence in the absence of the text. The examinee must judge whether each sentence means the same as a sentence that appeared in the passage.

There are four types of test sentences: originals, paraphrases, meaning changes, and distractors. An original test sentence is an exact copy of a sentence that appeared in the passage. Paraphrase sentences are constructed by changing as many words as possible in an original sentence without altering the meaning of the sentence. Meaning change sentences are constructed by changing one or two words in the original sentence so that the meaning of the sentence is altered. The final type of test sentence is a distractor. Distractors are sentences that have the same syntactic structure as an original sentence and are consistent with the overall theme of the passage, but are unrelated in meaning to any sentence in the passage and cannot be inferred from any sentence in the passage. Original and paraphrase test sentences have the same meaning as a sentence in the passage, and meaning change and distractor sentences have different meanings from sentences in the passage.

An SVT test typically consists of a set of three to six 12-sentence passages, each of which is followed by a set of 12 or 16 test sentences. In a 12-sentence SVT test, three of each test sentence type are used. In a 16-sentence SVT test, four originals, paraphrases, and meaning changes are selected to represent the 12 sentences in the passage, and then four distractor sentences are developed. After the test sentences are constructed, they are randomly arranged in the test with the constraint that test sentences measuring the first half of the passage appear first in the test. The reason for this restriction is to prevent an examinee from receiving a test sentence that has just been read or listened to, which would increase the chance that the examinee could respond to the test sentence based on the contents of short-term memory (Royer, 1990; Royer, Carlo, & Cisero, 1992).

In the present study, the SVT listening and reading comprehension tests each contained three 12-sentence passages and a 16-sentence test following each passage. The passages that comprised the listening and reading tests are modified versions of book reviews appearing in the Nonfiction in Brief section of the *New York Sunday Times Book Review* that have been used in previous research (e.g., Royer, Lynch, Hambleton, & Bulgareli, 1984; Royer, Marchant, Sinatra, & Lovejoy, 1990).

The SVT tests that follow each passage were modified for the present study. The original SVT tests that were used in previous research were 12-sentence tests (3 of each sentence type). The sets of test sentences were altered so that each set contained 16 test sentences instead of 12. In other words, one sentence of each sentence type was added to the original 12 to make four sentences of each sentence type for a total of 16.

Two forms of the SVT tests were constructed. Both Form A and Form B contained three listening and three reading passages. The listening passages in Form A were used as the reading passages in Form B, and the listening passages in Form B were the reading passages from Form A. A sample passage and test items are displayed in Appendix A.

CAAS Reading Battery

The CAAS battery is a computer-based system that measures the speed and accuracy of performance on tasks designed to assess component processes involved in reading (Royer & Sinatra, 1994). In each task, stimuli are presented by computer and examinees make responses into a microphone. The computer records the vocalization latency and the examiner records the accuracy of the response on-line by pushing a correct or incorrect button on a box connected to the computer.

The CAAS reading battery currently exists in three versions: an elementary version, a middle school version, and an adult version. The elementary version of the CAAS system consists of: a simple response time task, a letter naming task, word and pseudoword naming tasks, a category match task, and a semantics task. The middle school and adult versions of the CAAS system use word, pseudoword, category, and semantics tasks analogous to the elementary version except with more difficult stimuli. Examples of items in each of the elementary and adult tasks, which are used in the present study, are presented in Table 1 and a more complete description of each task is provided below.

Simple Response-Time Task. This task is a measure of the speed and accuracy of naming non-verbal stimuli. Examinees respond to displays of "***" or "+++" by saying "star" or "plus." The task is the first one presented in the battery and serves to acclimate the examinee to the testing situation.

Letter Naming Task. In this task, the examinee names an uppercase or lowercase letter that appears on the screen.

Word and Pseudoword Tasks. The word naming task requires examinees to pronounce single words. The words, which vary in length from three to six letters, have been reported by Dale and O'Rourke (1976) as familiar to at least 80% of fourth grade students.

The pseudoword task serves as a measure of phonological recoding ability, or the ability to apply grapheme-phoneme correspondences. Stimuli in the pseudoword task are pronounceable nonwords that have been derived from the real words by changing one letter in each real word.

Category Match Task. The category match task measures the ability to activate concepts in semantic memory. In this task, examinees are informed of the categories to be included in the task (transportation, animals, fruits, body parts, and clothes) and are then presented with pairs of words. Examinees indicate whether or not the words belong to the same category by saying "yes" or "no."

Semantics Task. This task assesses the application of semantic knowledge in sentence processing with a variation of the cloze procedure. Examinees are presented with sentences that contain a blank and a word above and below the blank. Subjects indicate which of the two words (which vary in semantic appropriateness) best fits the sentence.

Adult Word and Pseudoword Tasks. The adult word naming task consists of one-, two-, and three-syllable words with regular and irregular orthographic structure. Regular words are those with consonants and vowels having grapheme-phoneme correspondences that are relatively invariant across words. Irregular words are those that contain letters having grapheme-phoneme correspondences that are exceptions to the rule. For instance, "b" always take the \b\ sound except when it is silent in a small group of words such as "lamb" and "subtle." Therefore, a word like "bitter" would be considered regular and "debt" would be irregular. The criteria for regular and irregular grapheme-phoneme correspondences were taken from Venezky (1970). Half of the 1-, 2-, and 3-syllable regular and irregular words are low frequency (defined as less than 50 occurrences per million) and half high frequency (over 100 occurrences per million) (Francis & Kucera, 1982). Pseudowords were constructed from the real words by changing one letter per syllable.

Adult Category Task. This task is similar to the elementary category match task except that the categories are: politics, economy, and general science.

Adult Semantics Task. This task is also identical to the elementary level task, with the exception that the sentences are longer and the word choices are more complex vocabulary words (2- and 3-syllable regular and irregular words).

Table 1

<u>Examples of CAAS Tasks</u>

Task	Sample Stimuli
Simple	*** +++
Letter	A, g, K, n
Word	you goes horse banner
Pseudoword	yob poes porse danner
Category	YES: car/truck arm/leg NO: bus/stool nose/apple
Semantics	The farmer *planted*/played the corn.
Adult Word	sprint, plight, kitten, canoe, baritone, pseudonym
Adult Pseudoword	sprict, clight, fitken, yanob, larotine, psendinom
Adult Category	YES: delegation/ballot stock/bullish NO: voter/gene atoms/retail
Adult Semantics	A district attorney's job is to *prosecute*/perpetrate the defendant.

Phonological Processing Tasks. Three phonological processing tasks were included as part of the adult CAAS battery, and accuracy and response time data were collected. These tasks are visually presented "phonological awareness" tasks which measure an examinee's ability to detect rhyme, initial phonemes, and final phonemes in pairs of words. In each task there are four item types: (1) words that share the target sound and are orthographically similar, (2) words that share the target sound but are orthographically different, (3) words that do not share the target sound but are orthographically similar, and (4) words that do not share the target sound and are orthographically different. For instance in the rhyme task, two words that rhyme and are orthographically similar are *shoot/boot*, words that rhyme and are orthographically distinct are *shoot/fruit*, words that do not rhyme but are orthographically similar are *shoot/foot*, and words that neither rhyme nor share orthography are *shoot/walk*. All words in the three tasks are single syllable words that have been reported to be familiar to at least 80% of twelfth grade students (Dale & O'Rourke, 1976). Examples of the stimulus types for each task are displayed in Table 2.

Table 2

Examples of Stimuli in Phonological Processing Tasks

Task	Stimulus Pair
Rhyme	
Same Sound/Similar Orthography	pain main
Same Sound/Different Orthography	shoe two
Different Sound/Similar Orthography	food good
Different Sound And Orthography	trip late
Initial Phoneme	
Same Sound/Similar Orthography	chain chair
Same Sound/Different Orthography	phase flush
Different Sound/Similar Orthography	knit kite
Different Sound And Orthography	child open
Final Phoneme	
Same Sound/Similar Orthography	size doze
Same Sound/Different Orthography	trace lass
Different Sound/Similar Orthography	cheese chess
Different Sound And Orthography	niece splurge

Apparatus

A Sharp (Model 2000AV) tape recorder was used to present the SVT listening tests. Each passage was recorded followed by the 16 test sentences. Test sentences were separated by a 5 second interval. Reading comprehension tests were presented in test booklets along with general instructions for the comprehension tests and a sample passage and sample test items. A digital DECpc LPx 433dx desktop computer was used to administer all CAAS tasks.

Procedure

SVT listening and reading tests were group administered to 5 classes of nondisabled students (3 taught by 1 professor and 2 taught by another) during their introductory psychology class period. Nondisabled students from 2 classes taught by a third professor were individually administered SVT tests because the professor would not permit group administration of SVT tests to be conducted during class time. Learning disabled subjects were also individually administered the SVT tests. For all students, regardless of whether SVT tests were group- or individually-administered, the CAAS battery was individually administered at a subsequent testing session.

For the SVT tests, subjects were told that they would first hear three passages on a tape recorder, each of which would be followed by a set of test sentences, and then they would read three passages and respond to test sentences following each passage. They were then given instructions (which also appeared in the test booklet) about how to respond to test sentences. They were also given a sample passage and test sentences as practice. Administration of the listening and reading comprehension tests took approximately one hour.

For the CAAS battery, subjects were told that the object of each task is to respond as quickly and as accurately as possible. Tasks were presented in the following sequence: simple response time task, letter naming, word naming (elementary & adult level), pseudoword naming (elementary & adult level), phonological processing tasks (rhyme, initial phoneme, final phoneme), category match (elementary & adult level), and semantics (elementary & adult level) tasks. The reason for presenting the adult task after its corresponding elementary version was to decrease the amount of time needed for instruction since the task requirements for elementary and adult versions are similar. Administration of the above tasks took approximately one hour.

Data Cleaning

The data for all CAAS tasks are cleaned automatically by the CAAS program. The program eliminates responses faster than 250 milliseconds (since responses this short are impossibly fast). It then computes a mean and standard deviation for each examinee's set of responses, and trims the data by eliminating all responses more than two standard deviations above the mean. The computer program's final step is to recalculate the mean and standard deviation for each examinee's trimmed data.

CHAPTER 3

RESULTS

Reliability of Assessment Techniques

Reliability of SVT listening and reading tests (each of which contained 48 test items) was calculated using Cronbach's alpha (Cronbach, 1951). A reliability coefficient of .54 was obtained for the listening test and a coefficient of .40 was found for the reading test. The reliabilities were consistent with previous research findings that 48-item SVT tests have reliabilities between .5 and .6 (e.g., Royer et al., 1992).

Reliabilities on CAAS tasks could not be computed. One reason was that each subject received a different set of test items. For each task, a certain number of items was randomly sampled from a larger pool of test items so that each subject was given a slightly different test. A second reason was that the test length varied for different subjects due to the on-line deletion of items corresponding to microphone malfunction (microphone did not pick up subject's voice) or to on-line deletion of items where the microphone was activated by sources other than the subject's vocal response (e.g., coughing).

While it was not possible to calculate reliabilities of CAAS tasks used in the present study, there is documentation of the reliability of elementary-level CAAS tasks. A study by Sinatra and Royer (1993; Royer & Sinatra, 1994) with students in grades 2 through 5 used CAAS elementary reading tasks similar to those in the present study (except that in the present study more items were added to all tasks and all tasks used vocal responses rather than button responses). The authors reported reliabilities of response time measures on elementary CAAS tasks that ranged from .88 to .97 (Royer & Sinatra, 1994). Given that the elementary CAAS tasks in the present study were similar to the original tasks used by Sinatra and Royer (1993) and that the adult CAAS tasks were developed to be analogous to the elementary tasks, the reliabilities of tasks in the present study could be expected to be comparable to those obtained in the study by Sinatra and Royer (1993).[2]

Organization of the Chapter

There are several requirements that a diagnostic technique for identifying reading disability should satisfy. These requirements are used as a framework for presenting results regarding whether SVT and CAAS assessment techniques are useful for identifying reading disability in college students. The most basic requirement of an assessment technique that would be used for identifying reading disability would be to distinguish disabled from nondisabled students. Therefore, the first section presents results regarding whether SVT and CAAS can differentiate the nondisabled group from the learning disabled group.

An even more critical requirement of a reading diagnostic is that it identify reading disability and distinguish it from other types of disabilities. There are two ways to address this issue. One way would be to group students into broad diagnostic categories, such as reading disabled, generalized learning disabled, other types of disabilities, and nondisabled, and to determine whether the techniques can distinguish among the diagnostic groups. Results regarding this type of distinction are presented in the second section. The second way to address whether a reading diagnostic can distinguish reading disability from other disabilities is to classify disabled students on the basis of the problems they have (rather than classify students into broad diagnostic categories) and determine whether the techniques can distinguish disabled students with reading problems from disabled students with other problems. The third section presents results regarding whether the techniques can make this type of distinction.

The final requirement of a diagnostic technique is that it can be used to identify disabilities on an individual basis. The last section presents data on individual patterns of SVT and CAAS performance to address this issue.

Types of Analyses

The effectiveness of SVT and CAAS techniques, with respect to each of the above requirements of a diagnostic technique, was examined in two ways. First, the individual merits of SVT and CAAS were examined. Multivariate analyses of variance (MANOVAs) were used for evaluating each of these techniques by themselves. Second, the effectiveness of SVT and CAAS in combination was examined using discriminant analyses.

It is important to note why MANOVAs were chosen for evaluating the individual techniques rather than discriminant analyses. It would seem logical to use discriminant analyses to evaluate the techniques by themselves and in combination since results of the analyses could be compared to determine whether a discriminant function involving measures from one particular technique or from both techniques would best differentiate among the groups of interest. However, MANOVAs were chosen for two reasons. First, there was an interest in determining whether different patterns of performance on each of the techniques would be obtained for different groups of subjects (e.g., disabled versus nondisabled, or reading disabled versus other diagnostic categories).

Second, there are a number of problems with using discriminant analyses on the sample in the present study, and it was therefore decided that discriminant analysis should not be the primary tool with which to evaluate the assessment techniques. The first problem with using discriminant analysis in this case is that it requires known, predictable group membership. However, the actual grouping of subjects (e.g., disabled v. nondisabled) is uncertain. There may be some nondisabled subjects who actually have disabilities that have gone undetected, and conversely, there may be learning disabled students who actually have no disabilities. One alternative that does not require known group membership is cluster analysis. However, this is a more indirect approach and has its own set of limitations. A second problem with discriminant analysis in this instance is that analyses involving CAAS measures alone or a combination of SVT and CAAS measures include a large number of discriminating variables, and therefore, require much larger groups than those in the present study. A third problem is capitalization on chance. That is, there is the possibility of obtaining results, purely by chance, that indicate a "good" discriminant function. One way to circumvent this problem would be to randomly split the sample into two parts and conduct a discriminant analysis on each sample. However, large sample sizes are necessary to do this. Given the problems of using discriminant analyses with the present sample, it was decided that a more conservative approach would be to use MANOVAs wherever possible and to use discriminant analyses only when MANOVAs could not be used (i.e. when addressing whether techniques in combination could differentiate among groups).

Variables Included in Analyses

SVT administration resulted in 2 SVT variables, listening comprehension accuracy and reading comprehension accuracy. Proportion correct scores on the three listening passages were averaged to obtain a listening comprehension accuracy score, and proportion correct scores on the three reading passages were averaged to obtain a reading comprehension score. All analyses involving the SVT technique included a listening and reading score.

Administration of the CAAS elementary and adult reading batteries resulted in 26 CAAS variables (one accuracy and one response time score for each of the 13 tasks). Given the large number of variables, it was necessary to reduce the number of variables to as small a set as possible without sacrificing the wealth of information provided by the CAAS battery. Therefore, the reduction of variables involved combining data from similar tasks. Given that the elementary and adult versions of the word, pseudoword, category, and semantics tasks have similar task demands and assess the same cognitive processes (only at different levels of complexity), it would be reasonable to combine data from the elementary and adult tasks. Similarly, the three phonological processing tasks (rhyme, initial phoneme, and final phoneme) appear to tap a similar process, and therefore it would be reasonable to combine data from these tasks.

Support for combining tasks was provided by correlations which indicated a strong relationship between the elementary and adult tasks, and among the three phonological processing tasks. Correlations between elementary and adult task response times ranged from .58 to .88 ($p < .001$ for all correlations). Correlations between accuracy scores on the elementary and adult tasks were much lower (ranging from -.29 to .25), although this is most likely due to restriction of range given that performance was at ceiling on nearly every task. Correlations among the three phonological processing tasks ranged from .84 to .94 ($p < .001$ for all correlations) for response time data, and from .39 to .44 for accuracy data (again lower correlations may have been due to restriction of range).

Therefore, scores from the elementary and adult tasks were combined to form composite word, pseudoword, category, and semantics measures, and scores from the three phonological processing tasks were also combined to form a phonological composite. This was done separately for accuracy and response time measures. The end result was 7 CAAS measures of either accuracy or response time performance: simple, letter, composite word, composite pseudoword, composite category, composite semantics, and phonological composite. All MANOVAs on CAAS accuracy and response time data reported in the results section included these 7 variables.

With respect to discriminant analyses, if all SVT and CAAS measures were included, the number of discriminating variables would be 16 (2 SVT scores, 7 CAAS accuracy scores, and 7 CAAS response time scores). As noted earlier, a major concern was to keep the number of variables as small as possible (given the problems associated with small sample size). Therefore, the type of CAAS variables included in the discriminant analysis depended on the outcome of the MANOVAs on accuracy and response time data. If the MANOVAs performed separately on CAAS accuracy data and response time data revealed significant effects involving the grouping variable only for accuracy data, then only accuracy measures were included in the discriminant analysis with the SVT measures. If significant effects involving the grouping variable were only obtained in the response time analysis, then only response time measures were included in the discriminant analysis with SVT measures.

If both accuracy and response time analyses were significant, then combined accuracy/response time scores were used in the discriminant analysis rather than separate accuracy and response time measures in order to maintain a small number of discriminating variables. A description of how the combined accuracy/response time index (hereafter called the combined index) is calculated can be found in Appendix C. A final limitation of the discriminant analysis needs to be mentioned here, which is the possibility of capitalization on chance due to the fact that the discriminant analysis was not independent from the MANOVAs (since variables to be included in the discriminant analysis were chosen based on the outcome of the MANOVAs).

Analyses Examining Whether the Techniques Can Distinguish Disabled from Nondisabled Students

SVT

Table 3 displays the SVT performance of learning disabled and nondisabled students. A multivariate analysis of variance was performed on the listening and reading proportion correct scores with group (disabled v. nondisabled) as a between-subject factor and modality (listening v. reading) as a within-subject factor. A significant effect of group was found [F (1, 77) = 13.86, MSe = .01, p < .001] in that nondisabled students performed better than disabled students (74% for nondisabled as compared with 69% for disabled). There was also a significant effect of modality [F (1, 77) = 46.89, MSe = .004, p < .001] wherein performance on the reading test (75%) was generally better than on the listening test (68%). This overall pattern may reflect the fact that demands of the testing situation (e.g., reading passages are untimed and may be re-read, while listening passages are presented only once at a standard pace) lend themselves to better reading than listening performance (Royer et al., 1990).

Table 3

Proportion Correct Scores of Nondisabled and Learning Disabled Students on SVT Listening and Reading Tests

Group	SVT Test			
	Listening		Reading	
	Mean	Std Dev	Mean	Std Dev
Nondisabled	.71	.086	.77	.077
Learning Disabled	.65	.092	.73	.067

n=79

No significant interaction between group and modality was found [F (1, 77) = 1.53, MSe = .004, $p < .25$].

CAAS

Table 4 displays the accuracy and response time performance of nondisabled and learning disabled students on CAAS tasks. Separate MANOVAs were performed on accuracy and response time data with group[3] (disabled v. nondisabled) as a between-subject factor and task (simple, letter, word, pseudoword, category, semantics, and phonological) as a within-subject factor. The only significant result obtained from the accuracy analysis was an effect of task [F (6, 456) = 66.14, MSe = 19.5, $p < .001$] wherein accuracy slightly decreased as tasks became more complex. For the response time analysis, a significant effect of group was obtained [F (1, 76) = 26.02, MSe = .74], $p < .001$]. Nondisabled students were faster overall than disabled students. Task was also a significant source of variance [F (6, 456) = 357.98, MSe = .13, $p < .001$]. A significant group by task interaction was also found [F (6, 456) = 12.73, MSe = .13, $p < .001$]. A test of simple effects at each level of task (controlling for Type 1 error using the Scheffé procedure) indicated that disabled students were significantly slower than nondisabled students on all tasks except the letter task.

Table 4

Accuracy and Response Time (RT) Performance of Nondisabled and Learning Disabled Students on CAAS Tasks

Task	Group			
	Nondisabled		Learning Disabled	
	Mean	Std Dev	Mean	Std Dev
Simple Accuracy	98.9[a]	2.59	99.2	2.97
Simple RT	0.56[b]	.121	0.65	.219
Letter Accuracy	99.5	1.68	99.5	1.68
Letter RT	0.53	.083	0.56	.087
Word Accuracy	96.2	3.20	94.3	3.74
Word RT	0.63	.153	0.81	.260
Pseudoword Accuracy	91.2	7.47	88.7	8.26
Pseudoword RT	0.98	.615	1.43	.572
Category Accuracy	94.1	5.41	93.6	4.08
Category RT	1.39	.275	1.80	.453
Semantics Accuracy	93.2	6.62	93.2	5.40
Semantics RT	2.30	.515	3.03	.824
Phonological Accuracy	89.3	5.65	88.6	4.72
Phonological RT	1.49	.476	2.24	.910

n=78

[a] = percent correct

[b] = response time in seconds

SVT and CAAS Combined

A discriminant analysis was performed on group[4] (disabled v. nondisabled) by simultaneously entering all of the following variables: listening and reading proportion correct scores, and response time scores from the simple and letter tasks and from the composite word, pseudoword, category, semantics, and phonological tasks. Accuracy scores from the CAAS tasks were not included since the MANOVA on accuracy data (mentioned above) revealed no significant differences in performance among the groups.

The analysis revealed that 73.1% of the sample was correctly classified (where correct classification by chance would be 51%). Seven of the nondisabled subjects were misclassified as disabled. In interpreting the misclassification of nondisabled students, it is important to keep in mind the fact that the nondisabled sample may have contained students who had unidentified disabilities. Fourteen of the 36 disabled students (38.8%) were misclassified as nondisabled. Ten of these misclassified students had difficulties in areas other than reading. Therefore, it would seem logical that techniques which were designed to assess reading competence would assign students who have difficulties outside of reading a nondisabled status.

Taking into consideration the rate of correct classification by chance and the fact that the percentage of cases correctly classified by the discriminant function is an inflated estimate of the true percentage in the population (since the sample was used to derive the function and to test it), the obtained correct classification rate of 73.1% indicated that the discriminant function was performing relatively poorly. Other statistics which are indicative of a "good" discriminant function were also quite poor. The eigenvalue, which is the ratio of between-groups to within-groups sums of squares, should be large for "good" discriminant functions since good functions maximize the amount of between-groups variability. The eigenvalue of .53 obtained in the discriminant analysis, however, was relatively small. Moreover, the canonical correlation, which is a measure of the relationship between the discriminant scores and the groups, was rather low (.59).

The results of the analyses, taken together, suggest that disabled and nondisabled students do differ in their SVT and CAAS performance. Disabled students in general have poorer listening and reading comprehension than nondisabled students and perform more slowly than nondisabled on most CAAS tasks. Results of the discriminant analysis examining how well the SVT and CAAS techniques in combination can distinguish disabled and nondisabled students were somewhat discouraging. One possibility may be that a disabled versus nondisabled comparison may not be the best comparison since there are students with various types of disabilities within the disabled group. This variation within the disabled group may make it difficult to make an overall disabled-nondisabled distinction. A better comparison may be one which makes a finer distinction among disabilities within the disabled group. The next section presents results addressing whether the techniques can differentiate among different diagnostic categories.

Analyses Evaluating Whether the Techniques Can Differentiate Among Diagnostic Categories

This section addresses whether SVT and CAAS techniques could distinguish nondisabled students and different groups of disabled students. Recall from the Method section that students in the learning disabled sample were classified into 1 of 3 diagnostic categories based on clinical diagnosis and supplementary information: 1) reading disabled (RD), 2) generalized learning disabled (LD), and 3) other disabilities. Nondisabled students made up the nondisabled diagnostic category. The data presented in this section provide evidence for the effectiveness of SVT and CAAS techniques in differentiating nondisabled students, reading disabled students, generalized learning disabled students, and students with other disabilities.

Expectations for SVT Results

Nondisabled readers would be expected to perform better than all 3 disabled groups on listening and reading comprehension. Moreover, different patterns of listening and reading comprehension performance would be expected for RD and LD groups. The RD group would be characterized by average listening comprehension but poor reading comprehension. The reason for this pattern stems from the assumption of specificity that underlies definitions of specific reading disability. According to the assumption of specificity, modular processes rather than higher-level, global processes are areas of deficit in reading disability. Given that listening comprehension is a global process, poor listening comprehension would not be a key feature of specific reading disability. Reading comprehension, however, would be deficient since comprehension is likely to breakdown due to the poor word recognition skills of disabled readers.

The predicted pattern of performance for the LD group would be performance that is below average on both listening and reading comprehension. The reason is that students with generalized learning disabilities are most likely also poor readers. Therefore, poor readers, like RD students, would have reading comprehension problems. Poor readers, though, would also have listening comprehension difficulties given that they are characterized as having general cognitive deficits.

<u>SVT Results</u>

SVT performance of students in the nondisabled, reading disabled, learning disabled, and "other" diagnostic categories is shown in Table 5. A multivariate analysis of variance was performed on the listening and reading proportion correct scores with diagnostic category (RD, LD, other, nondisabled) as a between-subject factor and modality (listening v. reading) as a within-subject factor. A significant effect of diagnostic category was found [F (3, 72) = 4.34, MSe = .01, p < .01]. A set of planned contrasts comparing each of the three disabled groups to the nondisabled group (to control for Type 1 error, the Bonferroni inequality was used to set alpha at .017) revealed that the "other" group was significantly poorer than the nondisabled group (68% correct for the "other" group as compared to 74% correct for the nondisabled group) [t (72) = -3.06, SE = .028]. As in the analysis comparing the performance of disabled and nondisabled students, a significant effect of modality was found [F (1, 72) = 31.69, MSe = .004, p < .001] in which reading performance was better overall than listening performance (75% as compared to 68%). The interaction between diagnostic category and modality was not significant [F (3, 72) = 1.44, MSe = .004, p < .25].

Although the interaction was not significant, it is worthwhile to mention that students in the different diagnostic categories exhibited different patterns of listening and reading performance. Figure 5 displays the differential patterns of performance in terms of Z scores

Table 5

Proportion Correct Scores of Students in Different Diagnostic Categories on SVT Listening and Reading Tests

Diagnostic Category	SVT Test			
	Listening		Reading	
	Mean	Std Dev	Mean	Std Dev
Nondisabled	.71	.086	.77	.077
Reading Disabled	.68	.103	.71	.053
Learning Disabled	.63	.101	.74	.069
Other	.65	.092	.73	.067

n=76

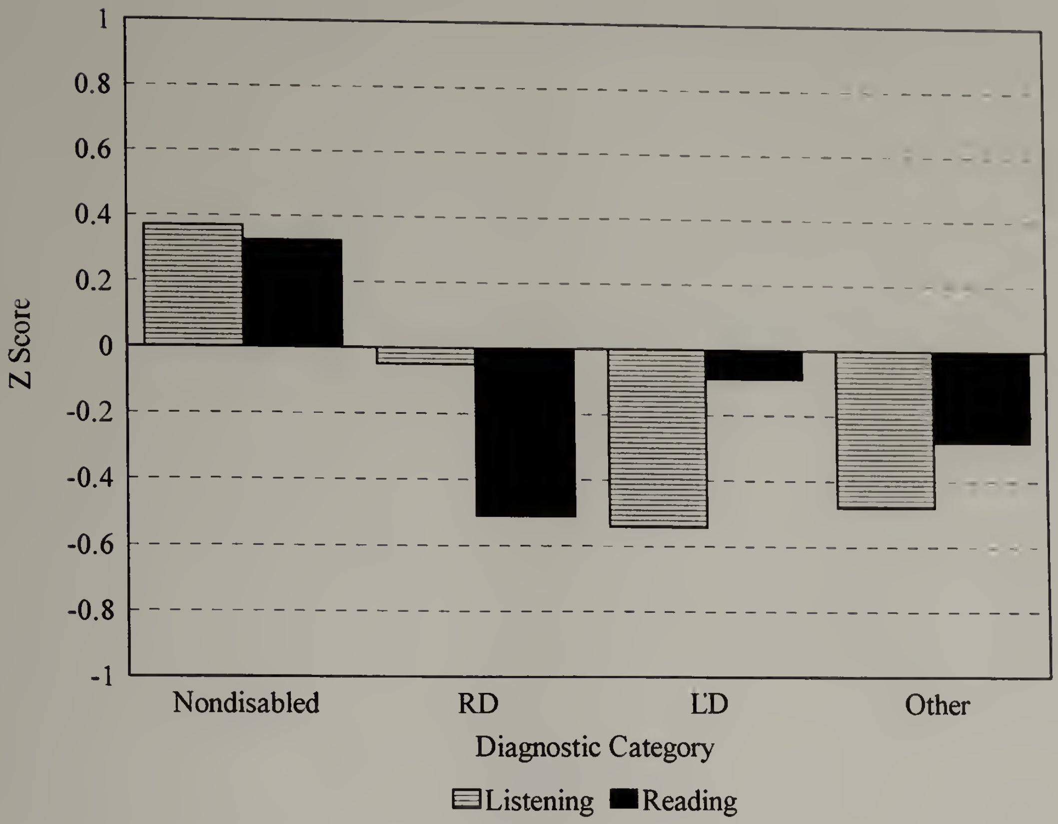

Figure 5. Z score performance of students in nondisabled, RD, LD, and "Other" diagnostic categories on SVT listening and reading comprehension tests.

which were converted from proportion correct scores for ease of interpretation. Z scores were calculated by subtracting the mean proportion correct scores of the entire sample from the mean score of each diagnostic category and dividing by the standard deviation of the sample. As can be seen in Figure 5, the performance of the nondisabled group was above the mean on both listening and reading, while the performance of the other diagnostic groups was below the mean. Moreover, the RD, LD, and "other" groups showed somewhat different patterns of listening and reading performance. Consistent with expectations, the RD group performed close to average on listening but below average on reading. The LD group showed the opposite pattern where listening performance was much poorer than reading. The poor listening comprehension of the LD group is consistent with the expectation that LD students would have difficulty in listening comprehension due to their general cognitive problems. The "other" group showed the same pattern as the LD group, except that the difference between listening and reading performance was less extreme.

Expectations for CAAS Results

Two hypotheses can be proposed on the basis of results from previous research on the CAAS system (e.g., Cisero et al., 1994). First, it would be expected that the overall accuracy and response time performance of both the RD and LD groups would be poorer than the nondisabled group. In contrast, the performance of the "other" group would be similar to that of nondisabled students. The reason is that students in this group have a specific disability outside of reading, and therefore their performance should be comparable to nondisabled students on a test of reading competence.

The second hypothesis is that different patterns of performance would be predicted for the RD and LD groups. The RD group would be expected to show performance that is comparable to nondisabled students on the simple and letter tasks, but much poorer performance on the rest of the CAAS reading tasks. The reason is that the word naming, pseudoword naming, and phonological tasks all involve phonological processing, which is hypothesized to be deficient in reading disabled students, and the category and semantics tasks involve word recognition, which is a skill that is affected by phonological processing.

In contrast, the LD group would be expected to have performance that is very poor across all tasks, including the simple task. The reason is that the LD group is characterized as having general deficiencies that cut across a variety of cognitive processes, which would affect performance on all CAAS reading tasks and on the simple task which has nothing to do with reading.

CAAS Results

Accuracy and response time performance on CAAS tasks of students in the four diagnostic categories is displayed in Table 6. Separate MANOVAs were performed on accuracy and response time data with diagnostic category[5] (RD, LD, other, nondisabled) as a between-subject factor and task (simple, letter, word, pseudoword, category, semantics, and phonological) as a within-subject factor. With respect to the accuracy analysis, diagnostic category was not a significant source of variance [F (3, 71) = 2.48, MSe = 51.08, p < .10]. In contrast, the response time analysis revealed a significant effect of diagnostic category [F (3, 71) = 22.71, MSe = .52, p < .001]. A set of planned contrasts comparing RD, LD, and "other" groups to nondisabled (to control for Type 1 error, the Bonferroni inequality was used to set alpha at .017) indicated that RD and LD groups were significantly slower overall than the nondisabled group [RD, t (71) = 3.47, SE = .293; LD, t (71) = 7.94, SE = .264].

The effect of task was significant in both the accuracy and the response time analysis [accuracy, F (6, 426) = 51.09, MSe = 18.74, p < .001; response time, F (6, 426) = 318.67, MSe = .11, p < .001]. Accuracy on the pseudoword and phonological composites was slightly lower than on the other tasks, and response time increased with the complexity of the task.

Table 6

Accuracy and Response Time (RT) Performance on CAAS Tasks of Students in Different Diagnostic Categories

Task	Diagnostic Category			
	Nondisabled	RD	LD	Other
Simple ACC[a]	98.9 (2.59)	100.0 (0.00)	97.6 (5.33)	99.6 (1.73)
Simple RT[b]	.569 (.121)	.578 (.109)	.822 (.364)	.604 (.092)
Letter ACC	99.5 (1.68)	99.2 (2.22)	100.0 (0.00)	99.7 (1.43)
Letter RT	.526 (.083)	.524 (.076)	.623 (.067)	.549 (.09)
Word ACC	96.2 (3.20)	93.1 (3.59)	91.9 (4.19)	96.4 (2.45)
Word RT	.633 (.153)	.799 (.112)	1.07 (.328)	.663 (.127)
Pseudoword ACC	91.2 (7.47)	82:3 (12.32)	87.2 (5.85)	93.2 (4.88)
Pseudoword RT	.975 (.615)	1.66 (.469)	1.78 (.494)	1.14 (.492)
Category ACC	94.1 (5.41)	93.8 (3.08)	93.8 (4.21)	93.3 (4.37)
Category RT	1.39 (.275)	1.87 (.425)	2.22 (.411)	1.56 (.351)
Semantics ACC	93.2 (6.62)	94.5 (4.11)	91.8 (5.03)	93.3 (6.22)
Semantics RT	2.30 (.515)	2.94 (.688)	3.74 (.646)	2.64 (.669)
Phonological ACC	89.3 (5.65)	84.5 (5.83)	88.8 (4.10)	90.4 (3.78)
Phonological RT	1.49 (.476)	2.20 (.540)	3.17 (1.19)	1.78 (.502)

n=75

[a] = accuracy (percent correct)

[b] = response time (seconds)

Standard deviations are in parentheses

A significant interaction between diagnostic category and task was obtained for the accuracy and response time analyses [accuracy, F (18, 426) = 2.27, MSe = 18.74, p < .01; response time, F (18, 426) = 8.61, MSe = .11, p < .001]. The nature of the interactions was that overall performance of the "other" group was comparable to the nondisabled group, but that the performance of RD and LD groups was poorer than the nondisabled group. Moreover, RD and LD groups showed different patterns of performance relative to nondisabled.

The differential patterns of performance of the diagnostic groups are depicted in Figure 6 in terms of percentile performance of RD, LD, and "other" groups as compared to nondisabled performance (represented as the 50th percentile). Some explanation of how these percentile scores were derived is necessary before discussing the results. First, accuracy and response time scores of the subjects on each task were combined into a single index of performance (called the combined index) using a transformation procedure developed by Sinatra and Royer (1995) that is explained in Appendix C.

The combined indices of performance of the RD, LD, and "other" groups were then transformed into effect sizes. Displaying the results in this way eliminates scale differences between the tasks. Effect sizes are calculated by subtracting the mean (here, the combined index) of the

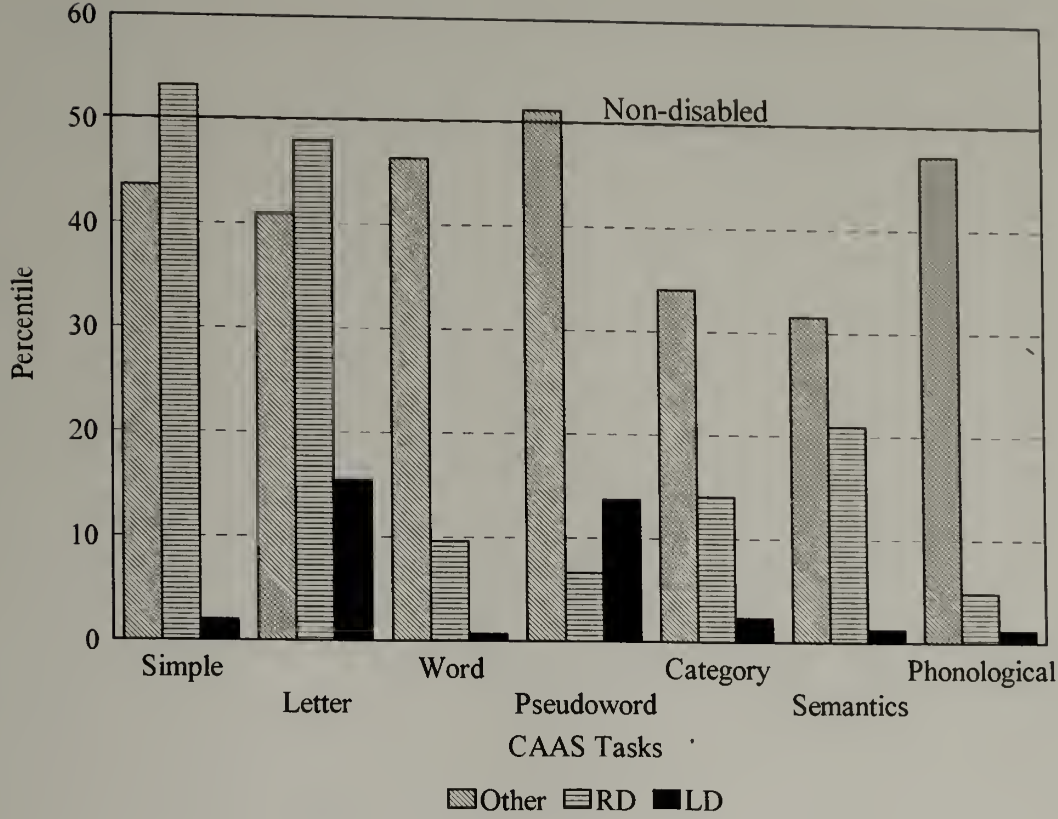

Figure 6. Percentile performance of RD, LD, and "Other" diagnostic categories on CAAS tasks as compared to nondisabled students. Nondisabled students are represented by the solid line at the 50th percentile.

experimental group (RD, LD, other) from the mean of the control group (nondisabled) and dividing by the standard deviation of the control group. The result is a Z score which indicates where the average experimental group subject would score _if_ he or she were in the control group. For further clarity of presentation, the effect sizes were transformed into percentiles.

As shown in Figure 6, the patterns of performance of the RD, LD, and "other" groups were consistent with the expectations. As predicted, the RD group performed as well as nondisabled on the simple and letter tasks, but was substantially worse on all other reading tasks. Also consistent with expectations, the LD group performed considerably worse than nondisabled across _all_ tasks (at or below the 10th percentile), even on the simple task which has nothing to do with reading.

Performance of the "other" group, as predicted, was comparable to nondisabled on most tasks, except for a slight drop in performance on the category and semantics tasks. The reason for the drop in performance on these tasks may be due to the fact that 8 of the 17 (47%) subjects in the "other" group were diagnosed with Attention Deficit Disorder (ADD) or had documented general knowledge problems or difficulties grasping meanings and ideas. Students with ADD may have experienced fatigue or a loss of attention on the category and semantics tasks since they were the last to be presented in the battery. Also, students with general knowledge problems or difficulties grasping word meanings would find the category and semantics tasks difficult since they assess the ability to activate concepts in isolation and in sentence contexts.

SVT and CAAS Combined

A discriminant analysis was performed on diagnostic category[6] (RD, LD, other, nondisabled) by simultaneously entering all of the following variables: listening and reading proportion correct scores, and combined (accuracy/response time) indices from the simple and letter tasks, and the word, pseudoword, category, and semantics, and phonological composites. The analysis revealed that the discriminant function performed relatively well. An eigenvalue of 1.54 and a canonical correlation of .78 were obtained. The word, phonological, semantics, and category composites made the best contributions to the prediction of group membership (correlations between the variables and the discriminant function ranging from .49 to .67).

The overall correct classification rate, however, was relatively low (70.7% as compared to chance classification of 42.9%). This was due mainly to the high number of "other" (12 of 17) subjects who were misclassified as nondisabled. Since all of the subjects in the "other" group were students who had a disability outside of reading, it would be expected that techniques designed to assess reading competence would assign these students a nondisabled status. There is a possibility, therefore, that there is no real difference on these measures between students with disabilities other than reading (i.e. the "other" group) and nondisabled students. If a distinction between "other" and nondisabled is not made, the overall correct classification would increase to 92%.

More important than the overall correct classification rate, however, is the ability of the discriminant function to identify the smaller groups of interest (Norušis, 1990). Since the purpose of this research was to determine whether SVT and CAAS techniques could identify reading disabled students as distinct from students with general cognitive deficits, the question of primary importance is whether the discriminant function can correctly distinguish RD and LD students. Therefore, the fact that the function correctly classified 5 of 7 (71.4%) RD subjects and 7 of 9 (77.8%) LD subjects is encouraging. Moreover, there were no RD subjects who were misclassified as LD, and conversely, there were no LD subjects who were misclassified as RD.

In sum, the results of the MANOVAs suggested that the SVT and CAAS techniques can distinguish among students in different diagnostic categories. Both SVT and CAAS data indicated different patterns of performance for students in the different diagnostic categories. Moreover, results of the discriminant analysis indicated that SVT and CAAS measures combined can distinguish reasonably well between RD and LD students, the two groups of primary interest. Therefore, it is possible that the differential patterns of SVT and CAAS performance obtained from the group data can be used in identifying individuals with reading disabilities and with generalized learning disabilities. Data examining the effectiveness of the techniques for identifying students on an individual basis will be presented in a subsequent section.

Analyses Evaluating Whether the Techniques Can Distinguish Among Different Types of Problems within the Disabled Group

As in the previous section, this section presents results indicating whether the techniques can identify reading disability and distinguish it from other types of disabilities. Rather than classifying students into broad diagnostic categories, the approach here was to group the disabled subjects into categories of problem types based on information that was available from the data sources listed in the Method section (i.e. self-report, description of difficulty from the official evaluation, standardized test scores, and history). One caveat is that the categories of problem types that were formed should not be considered to have known, predictable group membership. The reason is that classification was based on available information, much of which (as mentioned earlier) was missing, and if more information were available, some subjects may have fallen into different problem-type categories than those obtained.

SVT

An examination of the effectiveness of SVT in distinguishing among different problems within the disabled group involved evaluating whether performance on SVT differed for students with reported comprehension problems and students with no reported comprehension problems. Based on the information available for each student, subjects were classified as having comprehension problems (either listening, reading, or both problems)[7] or having no reported comprehension problems. Sixteen subjects were identified as having comprehension problems, and 21 were considered to have no comprehension problems.

It should be noted that students with no reported comprehension problems may have actually had comprehension problems that were not indicated by the available information. It is also noteworthy to mention that of students who were classified as having comprehension problems, few had standardized test scores to indicate this.

Table 7 displays the SVT performance of disabled students with comprehension problems and disabled students with no reported comprehension problems. A multivariate analysis of variance was performed on the listening and reading proportion correct scores with comprehension-problem group (comprehension problems v. no comprehension problems) as a between-subject factor and modality (listening v.

Table 7

Proportion Correct Scores on SVT Listening and Reading Tests of Disabled Students Who Have Comprehension Problems or No Comprehension Problems

Comprehension-Problem Group	SVT Test			
	Listening		Reading	
	Mean	Std Dev	Mean	Std Dev
Comprehension Problems	.61	.097	.71	.071
No Comprehension Problems	.67	.082	.75	.060

n=37

reading) as a within-subject factor. A significant effect of comprehension-problem group was found [F (1, 35) = 6.09, MSe = .01, p < .02] where students with reported comprehension problems generally performed more poorly than those with no reported problems (66% for those with comprehension problems as compared to 71% for those with no problems). There was also a significant effect of modality [F (1,35) = 27.84, MSe = .005, p < .001], as obtained in previous analyses where overall reading performance was superior to listening performance (73% as compared to 65%). No significant interaction between comprehension-problem group and modality was obtained [F (1, 35) = .31, MSe = .005, p < .60].

CAAS

An examination of the effectiveness of CAAS in distinguishing among different problems within the disabled group involved evaluating whether performance on CAAS differed for students experiencing difficulty in reading and students with other types of problems. Classification of subjects into a "reading problem" or "other problem" group was based primarily on the actual problems indicated by self-report, description of difficulty from the official evaluation report, test scores, and history. Subjects were classified as having a primary problem in reading if the available information indicated problems such as word recognition, decoding, reading comprehension, or slow reading rate. Subjects were categorized as having other problems if information indicated a problem unrelated to reading. Seventeen subjects were classified as having a reading problem (3 of whom had a reading plus another problem such as math) and 18 were classified as having other problems. Two subjects could not be classified due to inconsistent information from the various data sources.

Table 8 displays accuracy and response time performance of disabled students with reading problems and disabled students with other types of problems. Separate MANOVAs were performed on accuracy and response time data with problem-group[8] (reading v. other) as a between-subject factor and task (simple, letter, word, pseudoword, category, semantics, and phonological) as a within-subject factor. With respect to the accuracy analysis, task was the only significant source of variance [F (6, 192) = 33.3, MSe = 19.74, p < .01]. The response time analysis revealed a significant effect of problem-group [F (1, 32) = 4.33, MSe = .96, p < .05]. The group with reading problems was generally slower than the group with other problems. There was also a significant effect of task [F (6, 192) = 159.9, MSe = .18, p < .001]. The interaction between problem-group and task, however, was not significant [F (6, 192) = 1.64, MSe = .18, p < .15].

Table 8

Accuracy and Response Time (RT) Performance on CAAS Tasks of Disabled Students with Reading Problems and Disabled Students with Other Problems

Task	Problem Group			
	Reading		Other	
	Mean	Std Dev	Mean	Std Dev
Simple Accuracy	99.0[a]	3.85	99.6	1.68
Simple RT	0.69[b]	.308	0.61	.104
Letter Accuracy	99.2	2.07	99.7	1.39
Letter RT	0.59	.104	0.54	.068
Word Accuracy	93.1	4.46	94.9	2.75
Word RT	0.89	.319	0.75	.179
Pseudoword Accuracy	86.9	10.45	90.5	5.61
Pseudoword RT	1.60	.558	1.34	.566
Category Accuracy	93.5	4.46	93.4	3.89
Category RT	1.96	.509	1.65	.371
Semantics Accuracy	93.2	4.84	92.7	6.11
Semantics RT	3.29	.804	2.85	.781
Phonological Accuracy	88.3	5.67	88.8	4.18
Phonological RT	2.54	1.16	1.98	.564

n=34

[a] = percent correct

[b] = response time in seconds

SVT and CAAS Combined

In order to examine how well SVT and CAAS techniques in combination would be able to distinguish among disabled students with different types of problems, students needed to be grouped into categories of problem types that represented a range of problems that would be tapped by either technique. Therefore, students were grouped into the following categories of problems based on available information: 1) decoding only, 2) comprehension only, 3) decoding and comprehension problems, 4) reading plus another academic problem (e.g., math), and 5) other problems (i.e. problems outside of reading). One subject was categorized as having a decoding only problem, 8 as comprehension only, 5 as decoding and comprehension, 3 as reading plus another problem, and 18 as other. Two subjects could not be classified due to inconsistent information from the data sources.

A discriminant analysis was performed on the problem-type grouping[9] described above by simultaneously entering all of the following variables: listening and reading proportion correct scores, and combined (accuracy/response time) indices from the simple and letter tasks and from the composite word, pseudoword, category, semantics, and phonological tasks. Preliminary MANOVAs performed separately on CAAS accuracy and response time data indicated that problem-type group was a significant source of variance in both the accuracy and response time analyses. Therefore, it would be important to include accuracy and response time measures on CAAS tasks in the analysis. In order to reduce the number of discriminating variables, the combined indices were used in the discriminant analysis rather than separate accuracy and response time scores. To reiterate the caution stated earlier, membership of students in the problem-type categories is uncertain. This would affect the results of the discriminant analysis since discriminant analysis should be used for cases with known, predictable group membership.

The discriminant analysis produced an eigenvalue of 2.34 and a canonical correlation of .84. The simple and letter measures and the composite word measures were the best predictors of group membership (correlations between variables and the discriminant function ranging from .25 to .61). Moreover, the overall correct classification rate was 88.24% (where correct classification by chance would be 37.3%). All subjects in the "decoding only problems" or "reading plus other problem" groups were correctly classified. Also, seventeen of the 18 "other" subjects were correctly classified. Two subjects in the "comprehension only" group and 1 subject in the "decoding plus comprehension" group were misclassified as "other." It appears that groups characterized by comprehension problems were more difficult to classify. One reason may be that the CAAS measures were better predictors of group membership than SVT measures, and this may have affected the success of the discriminant function in classifying comprehension problems.

The results of the discriminant analysis indicated that the discriminant function predicted membership in problem-type categories relatively well. These results, while encouraging, need to be considered in light of the potential problems with using discriminant analysis in this instance (as discussed earlier in the results section). First, discriminant analysis requires that group membership is known. The actual grouping of subjects into problem-type categories was uncertain due to the paucity of information regarding their disabilities, and this would therefore reduce the effectiveness of the discriminant function. Second, there are problems related to the small sizes of the groups. One is that the large number of discriminating variables (9) requires much larger groups than those used in the present analysis. Another problem related to small sample size is the possibility of obtaining results that indicate a "good" discriminant function by chance alone. The small groups prevent splitting the sample into two halves and doing two discriminant analyses, thereby leaving open the possibility that results obtained from the single discriminant analysis could be due to chance.

The results of the MANOVAs and the discriminant analysis, taken together, suggest that SVT and CAAS techniques have the potential for differentiating among disabled students with different types of problems. Students who were identified as having comprehension problems did, in fact, perform significantly more poorly on SVT than those who had no reported comprehension problems. Students who were categorized as having reading problems were also significantly slower on CAAS measures overall than those with other types of problems. Moreover, the discriminant analysis suggested that a combination of SVT and CAAS measures can identify students with different types of problems.

Evidence on Whether the Techniques Can Be Used to Identify Disabilities in Individual Students

If SVT and CAAS are to be used for the purpose of identifying whether an individual student is reading disabled, then a given student's profile of performance on either of these techniques should indicate whether the student has a reading disability or some other type of disability. In order to evaluate the effectiveness of SVT and CAAS techniques for identifying the disabilities of individual students, individual patterns of performance on SVT and CAAS measures were examined to determine whether they were consistent with a student's diagnostic category. The use of SVT and CAAS measures in combination would also require that patterns of comprehension performance would be consistent with patterns of performance on the computer-based measure. Therefore, the degree of correspondence between SVT performance patterns and CAAS performance patterns was also examined.

Agreement between SVT Performance Pattern and Diagnostic Category

In order to determine the degree of correspondence between diagnostic category and patterns of listening and reading comprehension, it was necessary to group subjects into categories representing different patterns of listening and reading performance. Using a procedure identical to that of Carlisle and Felbinger (1991), subjects were grouped into 4 categories of SVT performance: 1) poor listeners and readers, 2) poor listeners/good readers, 3) good listeners/poor readers, and 4) good listeners and readers. Poor performance was defined as a score that was at least 1 standard deviation below the mean of the given subtest. It should be noted that according to this classification system, below average scores that were within 1 standard deviation of the mean would be considered "good" performance.

Expectations for Results. Nondisabled students were expected to be grouped into the good listeners/readers category. Most of the students in the "other" category, since they do not have reading disabilities, should also fall into this SVT performance group. The RD students would be predicted to fall into the good listeners/poor readers group. This pattern would be consistent with the assumption that the cognitive deficit of reading disabled individuals is specific to the reading process and does not affect general cognitive processes, of which listening comprehension is one. LD students would be expected to fall into poor listeners/readers or poor listeners/good readers groups since poor listening comprehension performance would indicate a deficit in general cognitive processes which characterizes students with general learning deficiencies. In contrast, RD students, by definition of a specific reading disability, would not be expected to show these patterns.

Results. Table 9 presents data indicating the number of subjects in each diagnostic category who were classified into each SVT performance group. The overall agreement between diagnostic category and SVT performance group was 64.5% (49 of 76 cases). Furthermore, a chi-square analysis performed on the data in Table 9 indicated no significant relationship between diagnostic category and the patterns of performance on SVT [X^2 (9, N = 76) = 10.29, p < .40].

One reason for the nonsignificant chi-square results and the low overall agreement between diagnostic category and SVT performance group may be that few of the disabled students fell into SVT performance groups that were consistent with their diagnostic category. Ten of the 17 (58.%) "other" subjects were classified, as expected, into

Table 9

Number of Subjects in Each Diagnostic Category Who Fall into Each SVT Performance Group

Diagnostic Category	SVT Performance Pattern			
	Good Listeners/ Readers	Poor Listeners/ Good Readers	Good Listeners/ Poor Readers	Poor Listeners/ Readers
Non-Disabled	35	4	2	1
RD	4	1	2	--
LD	8	1	--	1
Other	10	3	2	2

n=76

the good listeners/readers group. The remaining 7 subjects were about evenly distributed across the other 3 SVT performance groups. Only 2 of the 10 LD subjects fell into categories that represented poor listening comprehension performance. Moreover, only 2 of the 7 (28.6%) RD subjects fell into the expected good listeners/poor readers group. However, it should be emphasized that none of the RD subjects were classified as poor listeners/readers, which would be the pattern most at odds with a specific reading disability.

Closer examination of the mismatched RD, LD, and "other" students indicated that, in fact, many of them exhibited patterns of performance consistent with their disability. Three of the 4 RD students classified as good listeners/readers had better listening scores than reading scores, which is a pattern consistent with specific reading disability. Moreover, 4 of the 8 LD subjects in the good listeners/readers group had listening scores below the mean. Finally, 3 of the 7 "other" subjects who fell into unexpected SVT performance groups actually had patterns of performance that were consistent with their documented difficulties. For instance, the classification of an "other" subject into the poor listeners/readers group would be consistent with his/her documented difficulties in grasping meanings and ideas.

Agreement between Diagnostic Category and CAAS Profile Category

Previous research examining individual patterns of performance on CAAS tasks has indicated that students with reading problems show various types of performance profiles (e.g., Cisero et al., 1995, Wint, Cisero, & Royer, 1995). The profiles of 6 students are shown in Figures 7-12 as illustrations of different performance patterns. A brief description of each profile is given below.

As displayed in Figure 7, a profile where performance on all tasks, especially the simple task, is very poor as compared to other students has been termed the "global cognitive deficit profile" since poor performance is not limited to the reading tasks, but is also found on a non-verbal, perceptual task. In contrast, a profile which displays average to above-average performance on the simple and letter tasks relative to other students, but very poor performance on all other reading tasks (word and pseudoword naming, category, and semantics), especially word and pseudoword naming, relative to other students has been termed "specific reading disability" since the difficulty is localized in the reading tasks. Two types of specific reading disability profile have been found. The first, shown in Figure 8, is called "compensatory reading disability" because performance on simple, letter, and all (or most) elementary level reading tasks is average to

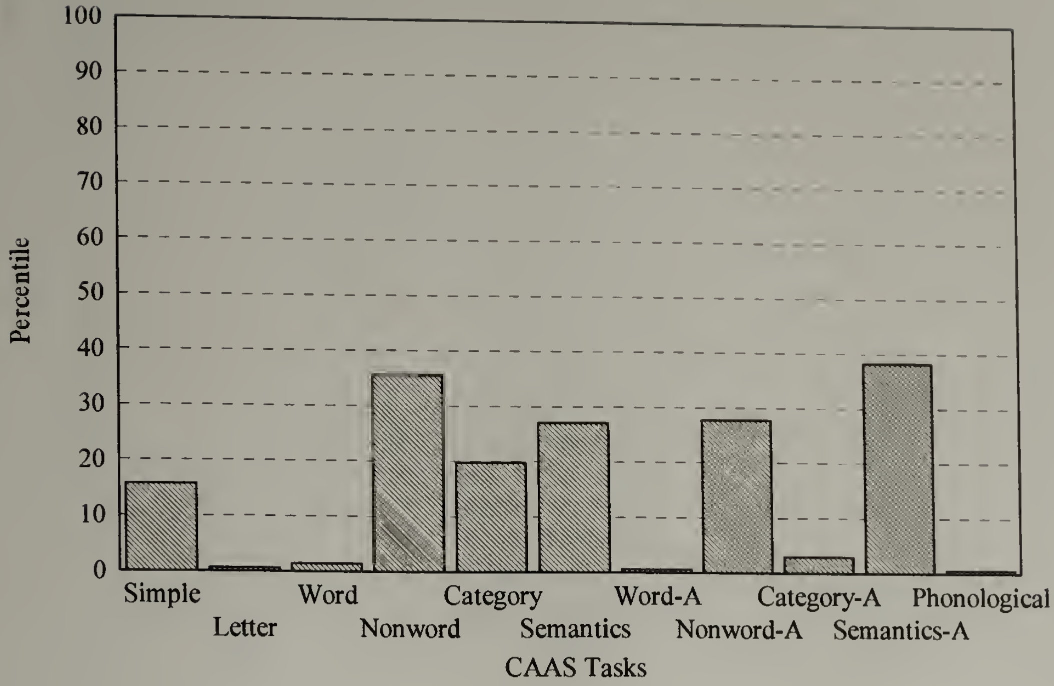

Figure 7. Prototypical profile of a global cognitive deficit.

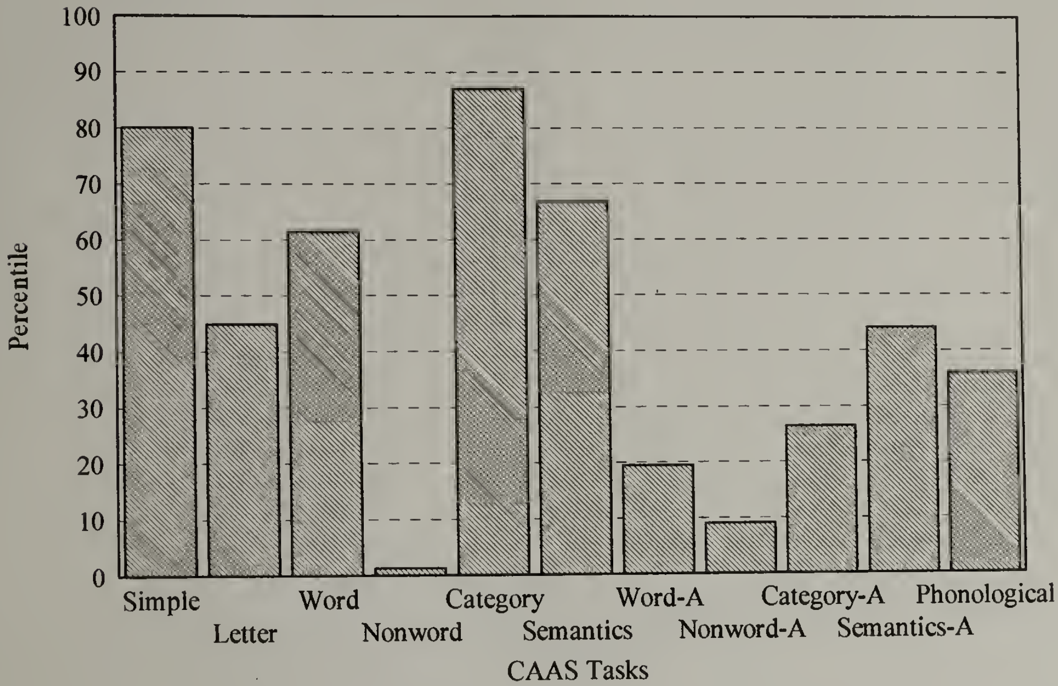

Figure 8. Prototypical profile of a compensatory reading disability.

above-average, but performance on all adult reading tasks is poor. The other profile shown in Figure 9, termed "severe reading disability," exhibits performance that is average to above-average on simple and letter tasks, but poor performance on all (or most) elementary level and adult level reading tasks. Finally, a "non reading disabled profile" has been found, which displays performance that is average to above-average on almost all tasks, as shown in Figure 10.

Two additional CAAS profile types have been found in the present sample that had not been present in previous research with smaller samples. Figure 11 displays a "meaning deficit profile," which is similar to a non reading disabled profile except that performance on category and semantics tasks (either at the elementary level, adult level, or both levels) is very poor. A "variable profile," as depicted in Figure 12, exhibits performance that appears to fluctuate according to the order of task presentation.

For the present sample, a CAAS profile for each subject was constructed by transforming his/her combined accuracy/response time index on each task into a percentile score. This was done by converting the combined index into a Z score and using the proportion of area under the normal curve corresponding to the Z score as the student's percentile rank. Composites of the elementary and adult level tasks were not used here (the phonological composite

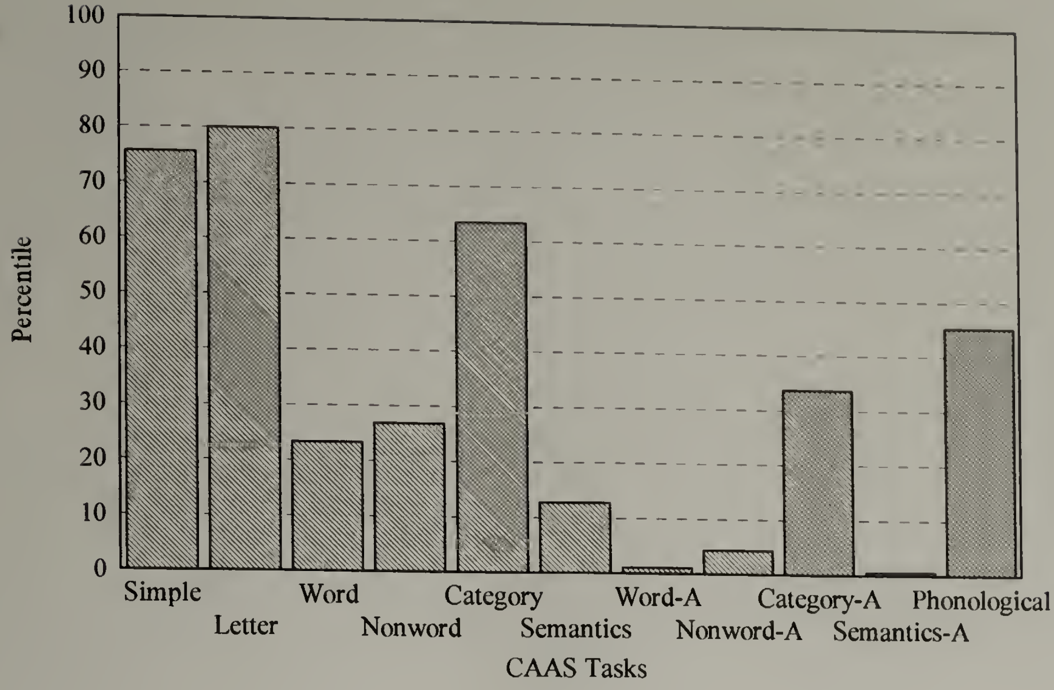

Figure 9. Prototypical profile of a severe reading disability.

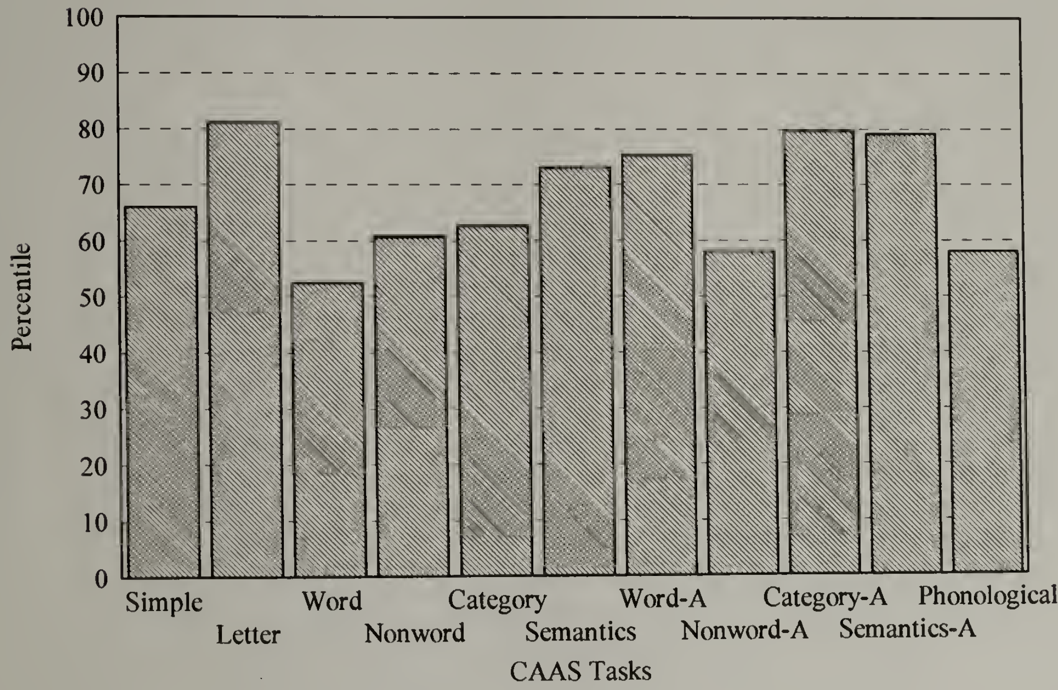

Figure 10. Prototypical non reading disability profile.

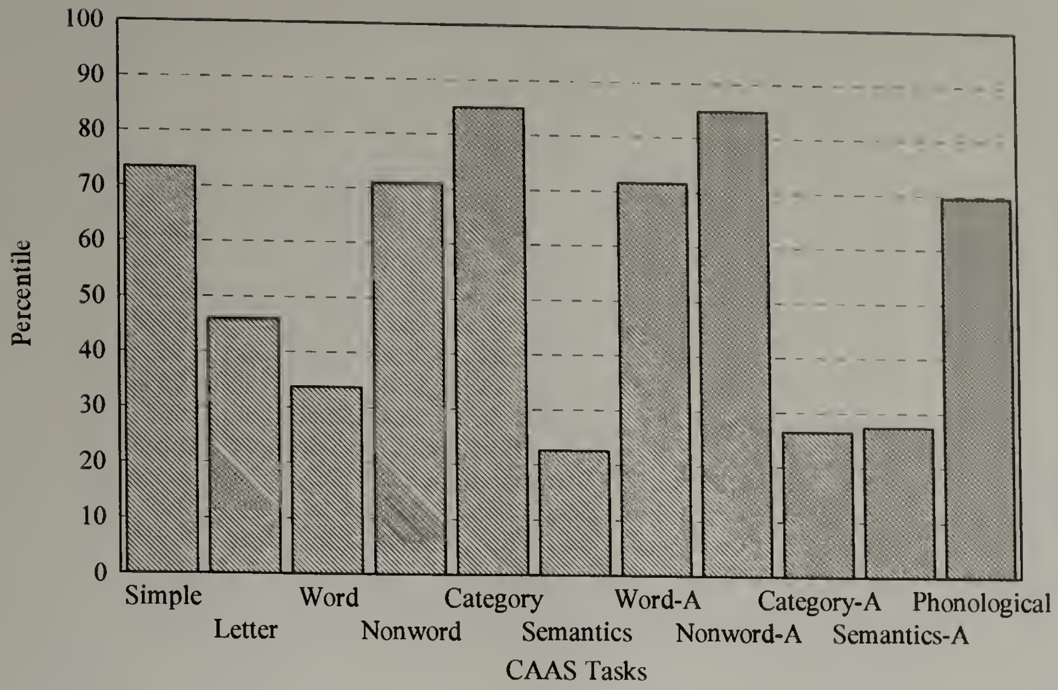

Figure 11. Prototypical profile of a meaning deficit.

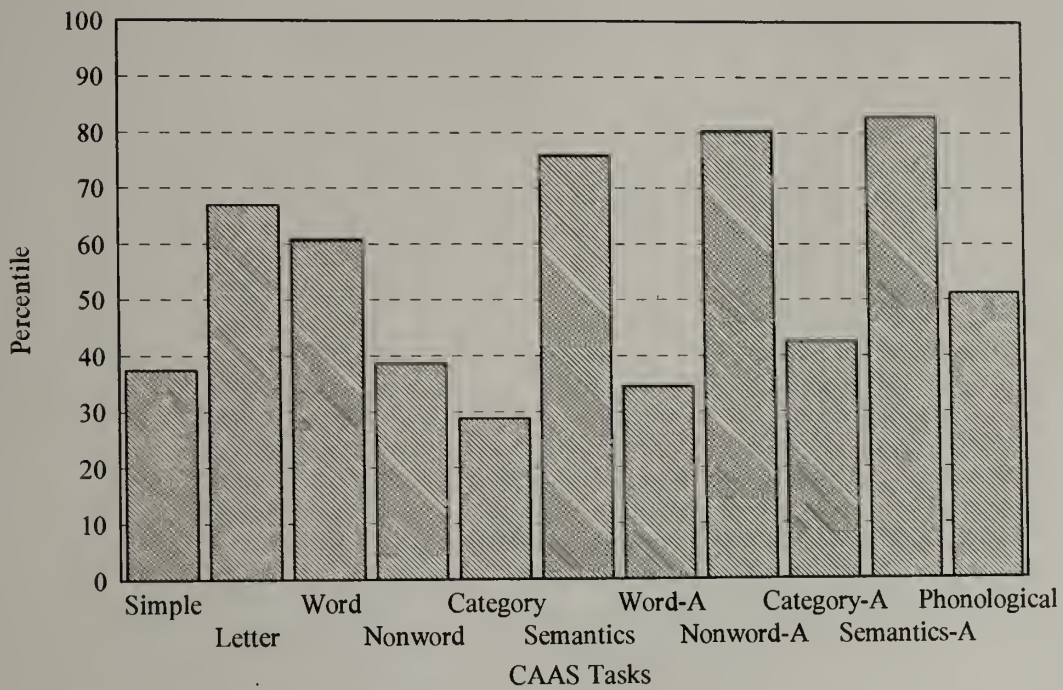

Figure 12. Prototypical variable profile.

was used, however) in order to preserve information regarding the two types of specific reading disability profile (severe and compensatory).

Using descriptions of CAAS profiles similar to the above description and prototypical CAAS profiles to represent each profile category, the author and a graduate student independently classified all subjects into CAAS profile categories. This was done without knowing the diagnostic category of the subjects. The overall agreement between the raters was 92.4% (where agreement by chance would be 37.4%). Therefore, in the interest of brevity, only the data using classifications of the author will be presented. The match between subjects' CAAS profile category and diagnostic category was examined.

<u>Expectations for Results</u>. Students in the RD diagnostic category should show a specific reading disability profile. In contrast, students in the LD diagnostic group would be expected to show a global cognitive deficit profile. The reason is that this diagnostic group is characterized as having generalized learning problems that cut across a variety of cognitive domains. Nondisabled students would be expected to exhibit a non reading disabled profile. Similarly, students in the "other" diagnostic category who have problems in math, long-term memory, and visual perception should show a non reading disabled profile. The reason is that these students have no

specific difficulty in reading. Students from the "other" diagnostic category who were identified as having problems related to general knowledge or grasping meanings and ideas should be expected to have a meaning deficit profile. Finally, students in the "other" category who have been diagnosed with Attention Deficit Disorder should exhibit a variable profile.

Results. Table 10 displays the number of students in each diagnostic category who showed each type of CAAS profile. A chi-square analysis performed on the data indicated a significant relationship between diagnostic category and CAAS profile category [X^2 (12, N = 76) = 32.91, $p < .01$]. The overall agreement between diagnostic category and CAAS profile category was 86.8%.

More important than the overall correspondence is how well diagnostic category and CAAS profile category match for the groups we wish to identify in practice, the RD and LD groups. Notice that all of the RD students were classified as having a specific reading disability profile, and all of the LD students were classified as having a global cognitive deficit profile.

Since it is unclear from the table how many students in the "other" category had consistent CAAS profiles, it is important to note that 13 of these students (76.5%) had CAAS profiles that fit with their documented difficulties. All 8 students in the non reading disabled profile category had

Table 10

Number of Subjects in Each Diagnostic Category Who Show Each Type of CAAS Profile

CAAS Profile Category	Diagnostic Category			
	Non-Disabled	RD	LD	Other
Non Reading Disabled	36	--	--	8
Specific Reading Disability	2	7	--	1
Global Cognitive Deficit	1	--	10	2
Meaning	2	--	--	4
Variable	1	--	--	2

n=76

difficulties outside of reading, all 4 in the meaning deficit profile category had documented problems in general knowledge or grasping meanings and ideas, and 1 of the 2 students who showed a variable profile had been diagnosed with Attention Deficit Disorder.

One possible explanation for the fact that 6 nondisabled subjects and 4 "other" subjects did not exhibit CAAS profiles consistent with their diagnostic category is that diagnostic category was not well-defined. That is, there may have been some subjects in the nondisabled category who actually had undetected problems that were revealed by the CAAS system. Likewise, the CAAS profiles of students in the "other" group may reflect difficulties that had not been previously detected by the testing that was done. For instance, the "other" student who exhibited a specific reading disability profile was diagnosed as having a math problem and there was no information on file to indicate that he or she had a reading problem.

Agreement between SVT Performance Pattern and CAAS Profile Category

Given the interest in whether CAAS and SVT information could be used together to identify reading disabled students, the correspondence between CAAS profile category and SVT performance pattern was examined.

Expectations for Results. Subjects who exhibit a "non reading disabled profile" would be expected to have good SVT reading comprehension. Subjects who show a "specific reading disability profile" should be classified as good listeners/poor readers. The predicted SVT performance groups for subjects who exhibit a "global cognitive deficit profile" would be poor listeners/readers or poor listeners/good readers since poor listening comprehension would be indicative of general cognitive problems.

Results. Table 11 displays the number of subjects (disabled and nondisabled combined) in each SVT performance group who exhibited each type of CAAS profile. A chi-square analysis performed on the data in Table 11 indicated a significant relationship between CAAS profile category and SVT performance group [X^2 (12, N = 79) = 31.81, $p < .01$].

As expected, most of the students who exhibited a non reading disabled CAAS profile (97.7%) had an SVT performance pattern that indicated good reading comprehension (good listeners/readers or good readers/poor listeners). However, fewer students showing a specific reading disability profile and global cognitive deficit profile fit their expected SVT performance patterns. Only 3 of 12 (25%) subjects with a specific reading disability profile fell into the expected good listeners/poor readers group. However, it should be noted that none of the students with a specific reading disability profile were classified as poor listeners/readers, the SVT pattern most at odds with a specific reading disability. Only 2 of 13 (15.4%) subjects with a

Table 11

Number of Subjects in Each CAAS Profile Category Who Show Each SVT Performance Pattern

CAAS Profile Category	SVT Performance Pattern			
	Good Listeners/ Readers	Poor Listeners/ Good Readers	Good Listeners/ Poor Readers	Poor Listeners/ Readers
Non Reading Disabled	38	5	1	--
Specific Reading Disability	8	1	3	--
Global Cognitive Deficit	11	1	--	1
Meaning	2	2	1	2
Variable	2	--	--	1

n=79

global cognitive deficit profile showed SVT performance patterns that could be consistent with general cognitive problems (poor listeners/readers or poor listeners/good readers).

A comment is needed regarding why the majority of subjects in the specific reading disability or global cognitive deficit profile categories were classified as good listeners/readers. The relatively good comprehension of these subjects may be due to the fact that SVT is an accuracy measure. Deficits such as specific reading problems or general cognitive problems, may only show up on speeded tasks (such as CAAS), especially at the college level where students with disabilities have acquired strategies to compensate for their difficulties that could be used on tests of higher-level skills such as SVT but that could not be used on speeded tasks.

CHAPTER 4

DISCUSSION

This research was undertaken to determine whether SVT and CAAS techniques could be used as alternatives to current diagnostic methods for identifying specific reading disability in college students. The search for new diagnostic techniques has become such an important issue due to the accumulation of evidence over the years which has indicated that current diagnostic procedures are inadequate for identifying reading disability. However, the inadequacy of current diagnostic methods makes the evaluation of new techniques difficult to say the least. The biggest shortcoming in the present research is the uncertainty in the criterion variable. That is, how can one be sure that the students who were receiving services from Disabled Student Services actually had learning disabilities or that the nondisabled control sample did not contain any students with undetected learning disabilities? Moreover, how could one be certain that the students identified as having a specific reading disability actually had reading problems? The answer is that one cannot be sure that all subjects in the nondisabled sample did not have any disabilities and all subjects in the disabled sample actually had disabilities, and further, that "reading disabled" subjects actually did have reading problems.

The best one could do would be to select students from the learning disabled sample that one could be reasonably certain had a specific reading disability. Classification of subjects was based primarily on the clinical judgment of the Counselor at Disabled Student Services (since this was the only source of information that was available for nearly every student). Four additional sources of data were collected in order to supplement the clinical judgment: self report of difficulty, description of difficulty from official evaluation report, standardized test scores, and history. Based on the clinical judgment and the supplementary data, students in the disabled sample were classified as reading disabled (RD), generalized learning disabled (LD), or "other."

Does Evidence Indicate That SVT and CAAS Techniques Can Identify Reading Disability?

The effectiveness of SVT and CAAS for identifying reading disability in college students was examined in four ways. First it was determined whether the techniques could differentiate disabled and nondisabled students in general. Next, and perhaps the more important question, was whether the techniques could distinguish reading disabled students from nondisabled students, from students with a generalized learning disability, and from students with other disabilities. Third, the ability of the techniques to differentiate among different types of problems within the disabled group was examined. Finally, the correspondence between SVT and CAAS data of individual subjects and their diagnostic category membership was examined to determine whether the techniques could be used to identify students on an individual basis. The results regarding each of these issues must be considered in light of the major limitation of the study, the uncertainty of group membership.

Disabled versus Nondisabled Distinction

A minimal requirement of diagnostic techniques for identifying reading disability is that they distinguish those with disabilities from nondisabled individuals. Separate SVT and CAAS analyses suggested that the techniques were able to differentiate disabled and nondisabled students. Evidence indicated that disabled students had significantly poorer SVT listening and reading comprehension than nondisabled students. Disabled students were also significantly slower than nondisabled students across CAAS tasks. However, results of a discriminant analysis using both SVT and CAAS measures indicated that the techniques in combination performed rather poorly in discriminating disabled and nondisabled. The reason may be that there was considerable variability within the disabled group with respect to the types of disabilities that subjects had, and that this variation made it difficult to make a broad discrimination between disabled and nondisabled using techniques that were specifically designed to detect difficulties in reading. This possibility is supported by the fact that most of the disabled subjects who were misclassified as nondisabled actually had disabilities in areas other than reading. Therefore, it may be that techniques designed to assess reading competence would be more successful at distinguishing reading disabled students from nondisabled students and from students with other types of disabilities.

<u>Distinction Among Different Diagnostic Groups</u>

A critical characteristic of diagnostic techniques for identifying reading disability is that they distinguish reading disabled students from nondisabled students and from students with other types of disabilities. Separate SVT and CAAS analyses revealed that students classified as reading disabled, generalized learning disabled, and students with other disabilities performed differently than nondisabled students on these measures and that each group showed a different pattern of performance.

With respect to SVT performance, nondisabled students had significantly better comprehension overall than the reading disability, learning disability, and "other" disability groups combined. The diagnostic groups also exhibited somewhat different patterns of listening and reading comprehension. Nondisabled students scored above the mean on both listening and reading comprehension, and the three disabled diagnostic groups (RD, LD, and "other") scored below the mean on both. Consistent with the notion of a specific reading disability, reading disabled students showed about average listening comprehension performance but poorer reading comprehension. The LD and "other" groups both had better reading than listening performance, with the difference between listening and reading performance being more extreme for the LD group.

With respect to CAAS performance, it was found that nondisabled students were significantly faster overall than the reading disability, learning disability, and "other" disability groups. Moreover, different patterns of performance were found among the diagnostic groups in both the accuracy and response time data. The combined accuracy/response time data shown in Figure 6 indicates that students in the "other" group performed similarly to nondisabled students on CAAS tasks. Students in the RD group performed as well as nondisabled students on the simple and letter tasks, but performed very poorly on all reading tasks. LD students, in contrast, performed very poorly relative to nondisabled students on all tasks, even on the simple task which has nothing to do with reading. The differential patterns of performance of the diagnostic groups that were obtained for both accuracy and response time data replicated a previous study which found distinct patterns of response time performance for RD and LD college students relative to a group of nondisabled college students (e.g., Cisero et al., 1994).

The distinct patterns of performance of the diagnostic groups on SVT and CAAS suggest that each of the techniques is able to differentiate nondisabled students and students with different types of disabilities. The discriminant analysis using both SVT and CAAS measures also indicated that the techniques in combination were successful in differentiating among the diagnostic groups. Therefore, there appears to be strong evidence that SVT and CAAS techniques, alone and in combination, are able to distinguish reading disabled students from nondisabled students and students with other disabilities, which is the purpose of a reading diagnostic.

There was evidence, however, that SVT was not as good as CAAS at distinguishing the diagnostic groups. First, the discriminant analysis revealed that CAAS tasks contributed more to the prediction of group membership than SVT measures. Second, different patterns of listening and reading comprehension performance were obtained for the different diagnostic groups (nondisabled, RD, LD, and "other"), but the interaction between modality and diagnostic category in the SVT analysis indicating the differential performance of the diagnostic groups was not significant. Reasons why SVT appeared to be less successful than CAAS as a diagnostic technique will be discussed in a subsequent section.

Distinguishing Reading Problems from Other Problems within the Disabled Group

The previous section discussed results regarding whether SVT and CAAS techniques could distinguish among students in different diagnostic categories. Another test of a reading diagnostic would be to classify disabled subjects according to the actual problems they have, rather than classifying on the basis of broad diagnostic categories, and to determine whether the technique can differentiate students who have reading problems from students who have other problems. Results indicated that disabled students with different types of problems performed differently on SVT and CAAS measures. Students identified as having comprehension problems were significantly poorer on listening and reading comprehension than those with no reported comprehension problems. Students identified as having a reading problem (e.g., word identification, decoding, slow reading rate) were significantly slower overall on CAAS tasks than students who were identified with other problems.

Moreover, a discriminant analysis using both SVT and CAAS measures correctly classified about 88% of disabled subjects into problem-type categories (decoding only, comprehension only, decoding and comprehension, reading plus another problem, and problems other than reading). All disabled subjects who were identified as having decoding problems or reading plus another problem were correctly classified. Also, the discriminant function correctly classified 94.4% of subjects who were identified as having problems other than reading and 75% of subjects who were identified as having comprehension problems only or decoding and comprehension problems. Again, as in previous discriminant analyses, CAAS tasks contributed more to the prediction of group membership than SVT measures.

Individual Patterns of Performance

The finding from group data that nondisabled students and students with different disabilities exhibited different patterns of performance on SVT and CAAS techniques is important in two respects. First, as discussed in an earlier section, it indicates that the techniques can be used to identify reading disability. Second, it suggests the possibility that the distinct patterns of performance of each diagnostic category found in the group data may be useful as a way of identifying individual students.

To explore this possibility, individual subjects were grouped into categories representing distinct patterns of performance on SVT and CAAS measures, and the match between the categories of performance and diagnostic category was examined. With respect to SVT performance, subjects were grouped into 4 categories of performance that have been used in previous research (e.g., Carlisle & Felbinger, 1991): 1) poor listeners and readers, 2) poor listeners/good readers, 3) good listeners/poor readers, and 4) good listeners/readers. The overall agreement between diagnostic category and SVT performance group was rather low (64.5%). Given that the majority of nondisabled subjects (83.3%) and most of the "other" subjects (58.8%) were classified, as expected, as good listeners/readers, much of the mismatch was found in the RD and LD groups. Most of the RD and LD students were classified as good listeners/readers, a pattern that is unexpected for both of these disability types. Closer examination of the data revealed that many of the RD and LD students who fell into this performance group actually had patterns of performance that were consistent with their diagnostic category (e.g., most RD students in the good listeners/readers group had better listening than reading scores).

More important, though, is the reason that RD and LD subjects fit the good listeners/readers classification in the first place. One possibility is that the "one standard deviation below the mean" cutoff for poor performance was too lenient. A more stringent cutoff, of say one-half standard deviation below the mean, may have resulted in correct classification of more RD and LD subjects. However, this would have been done at the expense of misclassifying more nondisabled and "other" students into categories of performance other than "good listeners/readers." Another more interesting possibility for why many RD and LD subjects fit the good listeners/readers classification is that SVT, being purely an accuracy measure, was not sensitive enough to detect the deficits of the RD and LD groups. It is very likely that disabled college students have acquired strategies to help them compensate for their disabilities that would enable them to perform relatively well on a test of comprehension.

In contrast to SVT, grouping of subjects into categories of performance on the CAAS technique proved much more successful. Subjects were classified as having one of five distinct CAAS profiles: specific reading disability, global cognitive deficit, non reading disabled, meaning deficit, and variable. The overall agreement between CAAS profile group and diagnostic category was about 87%. Moreover, all RD students were correctly classified as having a specific reading disability profile and all LD students as having a global cognitive deficit profile.

One reason that a better match was found between CAAS performance and diagnostic category than between SVT performance and diagnostic category is that the speed component of CAAS tasks is able to detect deficits characteristic of RD and LD students that would be undetected by accuracy measures such as SVT. This possibility is supported by the fact that many of the analyses of group data involving CAAS accuracy measures did not reveal significant effects.

In order to determine whether CAAS and SVT information could be used together to identify reading disabled individuals, the correspondence between CAAS profile category and SVT performance pattern was examined. Consistent with expectations, the majority of students showing a non reading disabled CAAS profile (about 98%) had good SVT reading comprehension performance. However, few students who had specific reading disability and global cognitive deficit profiles showed the expected SVT performance patterns. Rather, most students with a specific reading disability profile (67%) or global cognitive deficit profile (85%) were classified as good listeners/readers. The mismatch between CAAS profile category and SVT performance may be due to the fact that the CAAS technique is a speeded measure of performance while SVT is purely an accuracy measure. It may be that students in the global cognitive deficit profile or specific reading disability profile groups actually had general cognitive problems or specific reading deficits indicated by their CAAS performance. However, it is very likely that these students, by the time they have reached college, have acquired strategies to help them compensate for their disabilities. Compensatory strategies could be used on tests of higher-level skills such as comprehension, but the strategies would not be useful on speeded tasks that measure efficiency of performance on a given skill. As a result, the global cognitive deficits or specific reading deficits would only show up on speeded tasks. Therefore, students whose performance is poor on the CAAS measure appear to have relatively good comprehension on the SVT measure.

In sum, taking into consideration the major limitation of the study, which is the uncertain membership of students in the disabled and nondisabled samples, results regarding the usefulness of SVT and CAAS techniques for identifying reading disability in college students are encouraging. The techniques appear to be capable of distinguishing disabled from nondisabled, which is a basic requirement of a diagnostic technique for identifying reading disability. More importantly, SVT and CAAS techniques revealed different patterns of performance for nondisabled students, reading disabled students, and students with other types of disabilities, indicating that the techniques could be used to identify specific reading disability as distinct from other disabilities. Third, it was found that SVT and CAAS techniques reliably distinguished the performance of disabled students with reading problems and disabled students with other types of problems. Finally, the high degree of correspondence between CAAS profile categories and diagnostic category indicated that individual profiles of performance on the CAAS battery could be useful for identifying students who have a reading disability as well as students who have a generalized learning disability.

Does Evidence from SVT and CAAS Techniques Fit the Phonological-Core Variable-Difference Model?

One of the major advantages of using techniques such as SVT and CAAS as an alternative to the IQ-reading discrepancy is that they have been designed to tap the skills that are hypothesized to be deficient in disabled readers. The phonological-core variable-difference model proposed by Stanovich (1993) hypothesizes that the core deficit of individuals with reading problems is in phonological processing. Skills that involve phonological processing, such as phonological awareness, word identification, and decoding, would be deficient in individuals with specific reading disability and in poor readers. Reading comprehension, which would be affected by inefficient word identification processes, is another area of difficulty for both reading disabled individuals and poor readers. The term "variable-difference" refers to the fact that reading disabled students and poor readers would differ in areas outside of the phonological core. The deficits of disabled readers would be relatively specific to the phonological core, while poor readers would show a variety of cognitive deficits. Therefore, it would follow from the "variable-difference" argument that poor readers would have poor listening comprehension (as well as poor reading comprehension) since listening comprehension is a global cognitive skill, but that disabled readers would only have poor reading comprehension.

Based on the assumptions of the phonological core model, several hypotheses were proposed regarding the performance of reading disabled students on SVT and CAAS. The sections below discuss whether SVT and CAAS results fit the expectations.

SVT

It was hypothesized that reading disabled students would show a pattern of performance characterized by average listening performance but poorer reading performance. Results from the group data were consistent with the expectation. As shown in Figure 5, students in the reading disabled group had listening comprehension that was close to the mean performance of the entire sample but had somewhat poorer reading comprehension performance. Students from the LD and "other" groups showed the opposite pattern, and only the nondisabled group had both listening and reading performance that was above the mean. While Figure 5 shows different patterns of listening and reading comprehension performance for students in different diagnostic groups, a significant interaction between diagnostic category and modality was not obtained in the analysis of listening and reading proportion correct scores.

Support for the hypothesized pattern indicating normal listening comprehension and poorer reading comprehension was less clear in the individual data. Only 2 of 7 reading disabled students were classified into the expected SVT performance group, good listeners/poor readers. However, of the 4 reading disabled students misclassified as good listeners/readers, 3 exhibited better listening than reading performance.

<u>CAAS</u>

According to the phonological core model, reading disabled students would find difficulty in the word naming, pseudoword naming, and phonological processing tasks. Accuracy and response time analyses indicated that reading disabled students showed a pattern of performance consistent with this hypothesis. As shown in Figure 6, reading disabled students performed as well as nondisabled students on the simple and letter tasks, but performed very poorly relative to nondisabled on the word, pseudoword, and phonological tasks. Performance on the category and semantics tasks was also poor since concept activation and sentence processing, in part, require efficient word recognition processes.

It is noteworthy to contrast the performance of the reading disabled group with that of the LD group. The LD group, which consisted of students characterized by a variety of cognitive problems, may be considered poor readers since their global cognitive problems would affect reading as well as other domains. Notice that this group, like the reading disabled group, performs very poorly relative to nondisabled on tasks involving phonological processing. Also notice that in contrast to the reading disabled group, the LD group also performs poorly on tasks that tap processes outside the phonological core (i.e. the simple and letter tasks).

Two points need to be made about the distinct CAAS profiles of the RD and LD groups and how they relate to the phonological-core model. First, the phonological-core model is supported by the finding that both the reading disabled group and the learning disabled group (who may be considered poor readers) perform very poorly on tasks that tap phonological processing, but that only the LD group performs poorly on tasks that tap processes outside the phonological domain. Second, the fact that RD and LD students do not differ in the phonological core emphasizes the need for diagnostic techniques that can differentiate students with difficulties that are specific to the phonological core and students with difficulties that extend into other domains. The CAAS system has potential for making this distinction.

An examination of individual profiles revealed results that were consistent with the group data. All 7 reading disabled subjects showed a "specific reading disability profile" on CAAS tasks, which is characterized by average to above-average performance on simple and letter tasks but very poor performance on all reading tasks. Therefore, the individual data as well as group data indicated support for the notion that the deficits of reading disabled individuals are specific to phonological processing.

Why SVT Appeared Less Effective Than CAAS at Identifying Reading Disability

What becomes apparent from the above discussion of the effectiveness of SVT and CAAS techniques is that SVT appeared to be less successful than CAAS in a variety of respects. First, a significant difference in the patterns of listening and reading performance of different diagnostic groups was not obtained. Also, discriminant analyses using both SVT and CAAS variables indicated that CAAS measures were better predictors of group membership than SVT, regardless of what the grouping variable was. An examination of individual patterns of performance supported the group data. Most RD and LD students were classified as good listeners/readers, a pattern that is unexpected given their types of disability. Moreover, most students who exhibited a specific reading disability profile or global cognitive deficit profile on CAAS were also classified as good listeners/readers, which is a pattern of performance that is inconsistent with the types of deficits revealed by CAAS.

There are two possibilities to account for the apparent failure of SVT in reliably distinguishing reading disabled students from students with other types of disabilities. The first possibility, as mentioned in an earlier section, is that the accuracy-based nature of the test makes it insensitive in detecting reading disabilities at the college level. The SVT test measures an examinee's accuracy at answering test items that assess comprehension of the passage. Moreover, the reading portion of the test is untimed and an examinee is allowed to re-read the passage before answering test questions. The nature of the test, coupled with the fact that disabled college students have most likely acquired strategies to help them cope with their disabilities, may have made it difficult to find a large discrepancy between listening and reading comprehension in reading disabled students.

A second possibility for why the SVT measure appeared less effective than CAAS at differentiating reading disabilities from other disabilities and from no disabilities may be the poor reliability of the SVT test (reliabilities of .54 and .40 for listening and reading, respectively). Given that the magnitude of the reliability coefficient depends upon having variability in test scores and little error variability (Crocker & Algina, 1986), there are two possible factors that may have contributed to the low reliability of the SVT test in the present study.

First, low reliability may have been due to a small amount of test score variability within the sample. As indicated in Table 3, there was considerable overlap in the SVT performance of the 4 diagnostic groups. Moreover, Figure 5 indicates that although the diagnostic groups showed different patterns of listening and reading performance (as expected), each group had listening and reading performance that was within approximately one-half standard deviation of the overall mean.

A second factor contributing to the low reliability of the SVT test may be a considerable amount of error variability resulting from the fact that the test was a relatively small sample of a student's listening and reading comprehension ability. Lengthening the test would reduce the amount of error variability due to item sampling, thereby increasing the reliability. For instance, application of the Spearman Brown prophecy formula (Crocker & Algina, 1986) indicates that doubling the SVT listening test (to 6-passage tests having a total of 96 test items) would result in a reliability of .70 and doubling the reading test would result in a reliability of .57.

In sum, there is a possibility that an accuracy measure such as SVT may not be sensitive enough to detect reading difficulties of disabled students at the college level. Rather, it may be that the deficits of disabled college students are more easily detected by speeded tasks which put students in a situation where conscious strategies cannot be used to aid their performance. However, the possibility that the ineffectiveness of SVT in identifying reading disability is due to the accuracy-based nature of the test cannot be adequately examined until the reliability issue is resolved.

Advantages of SVT and CAAS Techniques over Current Diagnostic Procedures

Advocates of the status quo in learning disability diagnosis may argue, based upon the present results, that if the patterns of performance of students on SVT and CAAS techniques were consistent with their diagnostic category, then current diagnostic methods that are used to form the diagnostic groups would appear to be satisfactory. The response to this contention is that techniques such as SVT and CAAS have several advantages over current procedures which make them better diagnostic techniques.

First, present diagnostic procedures are very expensive and require an enormous amount of time and human resources. The IQ-achievement discrepancy described in an earlier chapter requires the administration of an IQ test and several standardized achievement tests in order to identify the particular disability of a student. A large battery of tests is needed, for instance, in order to properly determine that reading is the primary problem rather than mathematics, writing, and so forth. Moreover, information from standardized tests is often supplemented by other sources of data such as writing samples, interviews, and letters from high school counselors and teachers. Therefore, identification is almost never solely made by calculation of an IQ-achievement discrepancy, but frequently involves the clinical judgment of a single person or group of people based on information collected from various sources.

The diagnosis of a reading disability described above has several problems. First, the administration of a battery of tests by a trained professional is expensive, costing anywhere between $500 and $1200. Second, test administration requires a great deal of time and human resources. At one particular university, for instance, learning disability diagnosis requires up to 18 hours of assessments. Finally, and perhaps most importantly, extensive evaluations are not based on any theoretical models of the disability to be diagnosed (Morris, 1993).

The SVT and CAAS techniques, in contrast, require much less cost, time, and human resources. Individuals with little or no assessment experience can be trained to administer SVT and CAAS batteries, making these assessments very inexpensive. The total time for administration of both techniques is about two hours. Moreover, information obtained from these measures can be easily interpreted by a single person, even if the person has little familiarity with learning disabilities. Support for this is provided by the fact that the "naive" graduate student rater, who had no familiarity with the learning disability literature, was able to classify students into CAAS profile categories given only a prototypical profile and a brief description of the profile, and was generally very consistent with the classifications of the author.

A second advantage of SVT and CAAS techniques over current diagnostic procedures is that they assess the particular skills that are hypothesized to be deficient in disabled readers according to a model of reading disability. Current diagnostic procedures identify students as reading disabled if information from standardized tests and other sources indicate a reading problem and exclude difficulties in areas outside of reading. In contrast, diagnostic techniques based on a model of specific reading disability would allow a diagnostician greater precision in deciding whether or not a particular student is reading disabled. Patterns of performance that indicate the hypothesized deficits of disabled readers could be used as an indicator of specific reading disability.

A final advantage of SVT and CAAS techniques is their potential prescriptive value. As discussed in an earlier chapter, standardized reading achievement tests cannot adequately specify the particular deficits of disabled readers, and are therefore limited in terms of the prescriptions for remediation that they can provide. In contrast, results regarding individual patterns of performance from the present study and from a previous study (e.g., Cisero et al., 1995) suggest that profiles of performance on SVT and CAAS techniques provide information about the specific nature of disabled readers' problems. Information regarding the particular difficulties of a reading disabled student would allow a diagnostician to suggest instructional interventions that would help alleviate the reading problem.

Research presently being conducted at the Laboratory for the Assessment and Training of Academic Skills (LATAS) at the University of Massachusetts serves as an illustration of the potential usefulness of CAAS profiles for informing intervention. Children having various types of academic problems are brought to LATAS by parents who have exhausted all other avenues of help. Each student is given an initial testing battery consisting of SVT listening and reading comprehension tests and CAAS reading and mathematics batteries. Profiles of performance from the initial SVT and CAAS assessments are then used to determine the student's particular areas of difficulty and to determine the course of intervention.

Take, for instance, a third grade student who exhibits a severe reading disability profile (as shown in Figure 7c) where performance on the simple and letter tasks is average or above-average but performance on all other reading tasks is very poor relative to other students at the same grade level. This student would be considered to have a deficit in word recognition that serves as a bottleneck for the development of higher-level reading processes such as concept activation and sentence processing. Therefore, intervention begins by targeting the deficient skill, namely word recognition. LATAS uses automaticity training whereby children practice a given skill (e.g., word recognition) to improve the speed of their performance until their CAAS performance is comparable to grade-level peers. Once the deficient skill has been acquired to the point of proficiency and the bottleneck is removed, the child then moves to the next level in the hierarchy of reading tasks (in this case, concept activation) so that higher-level reading skills can then be developed.

SVT and CAAS techniques, therefore, have characteristics that make them more suitable diagnostic techniques than current procedures. They require less time to administer, and the theory-based nature of the techniques takes much of the guess-work out of diagnosis and prescription for remediation.

Future Questions

The findings from this research are encouraging as a first attempt at evaluating the usefulness of theory-based assessment techniques for identifying reading disability in college students. There are, however, unanswered questions that can be addressed in future research. One, for example, is: If the reliability of the SVT test used in the present study were improved, would SVT be as effective as CAAS as identifying reading disability in college students? The most important question, however, may be: Can we find a better way of classifying disabled subjects at the outset of the study (in order to eliminate the problem of uncertain group membership)? This appears to be the major obstacle for research on alternatives to the IQ-reading discrepancy approach for identifying reading disability.

APPENDICES

APPENDIX A

SAMPLE SVT PASSAGE AND TEST SENTENCES

Sample Passage

Mrs. Elizabeth: A Memoir was written by Elizabeth Anderson with help from Gerald R. Kelley. Mrs. Anderson, now eighty-four, was Sherwood Anderson's third wife. She met him in New York (where she was managing the Doubleday Doran bookstore) and lived with him in New Orleans, Paris, and rural Virginia until 1929. At that time, he sent her to visit her parents and then wrote her a one-line letter which stated: "I just wish you would not come back." Mrs. Anderson then moved to Mexico, renewed a friendship with William Spratling, whom she had known in New Orleans, and opened what became a successful dress shop. Her book ends with Spratling's death in an automobile accident in 1967, of which she comments: "I miss Bill Spratling so very much more than I ever missed Sherwood Anderson." It is a curious book, bland in describing her early years, dutiful and matter-of-fact about the Anderson years, and chatty about the Mexican years that followed. The writing is clearly that of Mr. Kelley, a professional journalist. But Mrs. Anderson's observations on her celebrated friends are just as clearly her own. "Others might eat an apple, Sherwood experienced it," she says. She also made such comments as "Edna St. Vincent Millay always had a coterie of followers but did not care about them one way or the other." Or, as she would observe of Bill Faulkner, "His studied courtesies and Southern mannerisms were a pose."

Sample Test Sentences

Original:	Mrs. Elizabeth: A Memoir was written by Elizabeth Anderson with help from Gerald R. Kelley.
Paraphrase:	The eighty-four year old Mrs. Anderson was the third woman to marry Sherwood Anderson.
Meaning Change:	She met him in New Orleans (where she was managing the Doubleday Doran bookstore) and lived with him in New York, Rome, and rural Pennsylvania until 1929.
Distractor:	For Elizabeth Anderson, Mrs. Elizabeth: A Memoir was her first book.

APPENDIX B

SUPPORT FOR THE RELIABILITY AND VALIDITY OF ASSESSMENT TECHNIQUES

A substantial amount of research has been conducted to document the reliability and validity of the SVT and the CAAS system. While there is no evidence available regarding the reliability and validity of the particular phonological processing tasks used in the present study, research indicates that phonological awareness tasks (which are similar to the tasks used in this study) are reliable and valid measures of phonological processing skill. A brief overview of the reliability and validity evidence related to the three techniques used in the study is presented below.

Reliability of SVT Tests

Royer and Hambleton (1983) report the reliability of SVT reading comprehension tests. The authors developed 50 passages at grade levels 3 through 7. The passages were then divided into 24 booklets, each of which contained 6 passages at adjacent reading levels (e.g., grades 4, 5, and 6 reading levels) and 16-sentence tests following each passage. The SVT tests were administered to over 1000 students. For each test booklet, a coefficient of internal consistency was calculated (based on 96 test sentences). The mean reliability was .92 with coefficients ranging from .84 to .98.

A study by Royer, Kulhavy, Lee, and Peterson (1986) reports evidence that the SVT is a reliable measure of both listening and reading comprehension. Grade 4 and 6 students were administered SVT listening and reading comprehension tests based on passages at grade 3, 5, and 7 readability levels. Both the listening and the reading test contained 3 passages and 16-sentence SVT tests, making a total of 48 test sentences for both the listening and reading SVT test. For the reading test and listening test, test sentences were divided to form two test scores, and a corrected split-half reliability coefficient was computed. A reliability of .85 was obtained for the reading test and .71 for the listening test.

Several other studies have reported reliabilities for SVT tests (e.g., Greene, Royer, & Anzalone, 1990; Royer & Carlo, 1991; Sinatra, 1989). In general, research has shown a relationship between the reliability of an SVT test and test length. SVT tests typically have reliabilities between .5 and .6 for tests based on three 12-sentence passages and 16-sentence tests (48 test sentences), between .7 to .8 for tests based on four passages (64 test sentences), and between .8 to .9 for tests based on six passages (96 test sentences) (Royer et al., 1992).

Validity of SVT Tests

Evidence for the validity of the SVT as a measure of comprehension has accumulated from numerous studies. Research has indicated that the SVT is sensitive to text difficulty, and to differences in reading skill, that SVT performance varies as a function of working memory capacity (which is a key factor in comprehension), that performance on SVT listening and reading tests is consistent with what theory states about the relationship between listening and reading comprehension, and that SVT shows good convergent and divergent properties. Each of these types of validity evidence is discussed briefly below. Royer (1990) provides a more detailed discussion of these types of evidence, as well as a discussion of other types of validity evidence.

Sensitivity of SVT to Text Difficulty. One characteristic that tests of comprehension should have is to be sensitive to the difficulty level of the text. Research has indicated that SVT is sensitive to text difficulty. For instance, studies which entailed administering SVT tests based on passages drawn from texts used in different grades have found that SVT performance declined as a function of text difficulty (Greene, Royer, & Anzalone, 1990; Royer & Carlo, 1991; Royer, Hastings, & Hook, 1979; Royer et al., 1986). Also, a study by Royer and Hambleton (1983), which involved administering SVT tests based on passages with readabilities ranging from grade 3 to grade 7, found that SVT performance systematically varied as a function of the readability of the passages.

Sensitivity of SVT to Differences in Reading Skill. Another quality of comprehension tests is that they should be sensitive to differences in reading skill. Studies by Royer et al. (1979, Experiment 2; 1986) have provided evidence that SVT performance varies as a function of grade level. Moreover, research has indicated that SVT is sensitive to differences in reading skill when reading competence is defined by external criteria. For instance, several studies have reported evidence that SVT performance varies as a function of teacher ratings of reading skill (Rasool & Royer, 1986; Royer & Carlo, 1991; Royer, Carlo, Carlisle, & Furman, 1991). Further, there is evidence that children with high scores on standardized reading test perform significantly better on SVT tests that children with lower standardized reading test scores (Royer, Sinatra, & Schumer, 1990).

SVT and Working Memory Capacity. Limited working memory capacity has been hypothesized as one source of reading comprehension difficulty. This hypothesis is based on the notion that working memory is necessary for holding a sufficient number of linguistic units until a meaningful unit can be accumulated and interpreted. Lynch (1986, 1987) has provided evidence that SVT performance is related to working memory capacity. Significant correlations of .59 (Lynch, 1986) and .67 (Lynch, 1987) were found between performance on a working memory task and SVT performance. Moreover, SVT performance differed significantly between students who performed poorly on a working memory task and those who performed well (Lynch, 1986).

Performance on SVT Listening and Reading Comprehension. One assumption that underlies most theories of comprehension is that competence in listening comprehension develops before reading comprehension and places an upper limit on reading comprehension performance (Royer, 1995). Research indicates that the relationship between listening and reading comprehension generally fits with this notion. In a study by Royer et al. (1986), which involved presenting students in grades 4 and 6 with SVT listening and reading comprehension tests based on passages with grade 3, 5, and 7 readabilities, reading comprehension exceeded listening comprehension on passages having readabilities below the grade level of the students, but listening was superior to reading on passages having readabilities above the grade level of the students. A study by Royer, Sinatra, and Schumer (1990), which administered SVT listening and reading tests to students in grades 3 and 4, revealed that good readers performed better on reading than on listening, while the opposite was true for poor readers. The results of these studies indicate that listening is superior to reading when the student has very poor reading skills or when materials exceed the capabilities of the student, and that reading is superior to listening when the student has well-developed reading skills or when materials are sufficiently easy.

Convergent and Divergent Validity Evidence. Research indicates that SVT tests are positively related to other measures of reading comprehension and are not strongly related to measures that do not depend on reading skill. SVT has been shown to correlate .5 with Stanford Achievement Test reading comprehension, .73 with Iowa Test of Educational Development reading comprehension (Royer et al., 1979), and .52 with California Achievement Test reading comprehension (Royer, 1995). Correlations between SVT and other measures requiring reading skill (e.g., reading achievement in science or social studies) range from .58 to .73 (Royer, 1990). In contrast, correlations between SVT and measures of mathematical computation and math concepts range from .15 to .28 (Royer, 1990).

Reliability of the CAAS System

Support for the reliability of the CAAS system can be found in a study by Sinatra and Royer (1993; Royer & Sinatra, 1994). Students in grades 2 through 5 were administered SVT listening and reading comprehension tests and the elementary reading battery of the CAAS system, and a subset of these children were readministered the CAAS battery one year later. At the time of this study, the elementary version of the CAAS system was comprised of a simple response time task (responding to "***" or "+++"), a variation of the Posner letter match task (e.g., Posner et al., 1969), word and pseudoword naming, a category match task (deciding if two words belong to the same category), and two variations of a cloze task designed to measure syntactic and semantic analysis of sentences (syntax and semantics tasks). Accuracy and response time were recorded on each task.

Reliability of response time measures was established by estimating components of variance attributable to subjects and to stimulus items for each of the tasks. Reliabilities on CAAS tasks ranged from .88 to .97 with a mean of .94 (Royer & Sinatra, 1994).

Validity of the CAAS System

The study by Sinatra and Royer (1993; Royer & Sinatra, 1994) described above evaluated the validity of the CAAS system by examining relationship between performance on CAAS tasks and various indices of reading skill. One index of reading skill was grade level since students in higher grades would presumably be better readers than younger students. Response accuracy, which averaged over 90% correct across the tasks, did not vary as a function of grade. The exception was a significant improvement in word and pseudoword naming as grade level increased. Response time on all tasks significantly decreased as grade level increased.

Another index of reading skill was the reading book level of students in grades 2 to 4. Word and pseudoword naming response time varied systematically with reading ability for grades 2 and 3. Response time on the semantics task (assessing sentence processing) significantly discriminated between ability levels in grades 2 to 4. Moreover, for grade 3 the category and syntax tasks discriminated among ability levels in addition to the semantics task.

Another type of validity evidence was to sort students into high, average, and poor groups based on CAAS performance using a cluster analysis method and to examine teacher ratings of reading competence for the groups. The authors found that teacher rating of reading competence at the one-year follow-up varied in accordance with clusters formed from the CAAS assessment one year earlier.

A final type of validity evidence for the CAAS system involves data demonstrating that CAAS performance is consistent with a cognitive-developmental theory of reading (see Royer & Sinatra, 1994 for a review). The authors found that word identification skills of grade 3 and 4 students were more strongly related to sentence processing than the word identification skills of grade 2 students. This finding is consistent with cognitive-developmental theory. Young readers' word identification is slow and not yet automatic. Much of their cognitive capacity is used for word identification so that little is left for comprehension. In contrast, older readers have developed more efficient word identification processes which allows capacity to be used for higher-level comprehension activities. Thus, for older readers word identification makes a stable and consistent contribution to sentence comprehension (Royer & Sinatra, 1994).

Reliability of Phonological Awareness Tasks

Despite the large amount of research on phonological awareness, evidence supporting the reliability of phonological awareness tasks is rarely reported. Studies by Stanovich, Cunningham, and Cramer (1984), Yopp (1988), and Cisero and Royer (1995) are a few exceptions.

In the study by Stanovich et al. (1984) kindergarten students were given 10 phonological awareness tasks: rhyme production, rhyme detection, detecting same initial consonants, detecting different initial consonants (done with instructions phrased in two ways), detecting same final consonants, detecting different final consonants, deleting initial consonants, substituting initial consonants, and isolating the initial consonant. Split-half reliabilities of the tasks ranged from .63 to .95 with a mean reliability of .81.

Similarly, the study by Yopp (1988) involved administering a battery of 10 phonological awareness tasks to kindergarten students. The tasks in Yopp's battery included: auditory discrimination, rhyme detection, phoneme blending, phoneme counting, two variations of a phoneme deletion task, two variations of a phoneme segmentation task, sound isolation (initial, medial, final), and word-to-word matching (detecting similarities in words). Reliabilities (Cronbach's alpha) ranged from .58 to .96 with seven of the ten tasks having reliabilities greater than .83.

Cisero and Royer (1995) used phonological awareness tasks that most closely resemble the tasks used in the present study. Native English-speaking and native Spanish-speaking kindergarten and first grade children were administered rhyme detection, initial phoneme detection, and final phoneme detection tasks in which pairs of 3-phoneme Consonant-Vowel-Consonant words were presented by tape recorder and students made button responses as to whether the two words had the same or different target sound. Kindergarten students received tasks only in their native language and first grade students were administered tasks in both language to examine transfer of phonological awareness skills. Reliability indices (Cronbach's alpha) for rhyme detection accuracy were .78 for English and .69 for Spanish. Reliabilities for initial phoneme accuracy were .51 for English and .71 for Spanish, and for final phoneme were .59 for English and .62 for Spanish.

Results from these studies, therefore, suggest that tasks assessing one's sensitivity to the phonological structure of spoken words have moderate to high reliability.

Validity of Phonological Awareness Tasks

A large body of evidence has accumulated which indicates the validity of phonological awareness tasks as measures of phonological processing skill that is critical for successful reading acquisition. If phonological awareness tasks are tapping a capability that is necessary for reading success, then two requirements of phonological awareness tasks would be that: 1) they predict beginning reading achievement, and 2) they differentiate successful readers from those with reading problems.

There is ample evidence to support both criteria. Evidence indicates that performance on phonological awareness tasks prior to formal instruction is predictive of reading achievement once instruction is begun (Bryant et al., 1990; Bryant Maclean, Bradley, & Crossland, 1990; Lundberg et al., 1988; Maclean et al., 1987; Perfetti et al., 1987). Research also suggests that disabled readers perform significantly more poorly on phonological awareness tasks than younger nondisabled readers of the same reading level (e.g., Bruck, 1992; Bruck & Treiman, 1990; Ellis & Large, 1987; Manis et al., 1988; Pennington et al., 1990).

Further evidence for the validity of phonological awareness tasks is provided by factor analyses which indicate that a considerable amount of variance is shared by various measures of phonological awareness. In the study by Stanovich et al. (1984) mentioned above, a principal factor analysis performed on the 10 phonological awareness tasks revealed that one factor accounted for 47.8% of the variance. Seven of the 10 tasks loaded highly on the factor and the remaining three tasks had low to moderate loadings. The study by Yopp (1988) found that all phonological awareness tasks loaded on one of two factors, which together accounted for 68% of the variance. Specifically, tests of phoneme blending, segmentation, and counting, and sound isolation tests loaded highly on Factor 1, and tests of phoneme deletion loaded highly on Factor 2. The word-to-word matching task had a moderate loading on Factor 2, while the auditory discrimination and rhyme tasks had low to moderate loadings on both factors.

The data, therefore, support the reliability and validity of the three assessment techniques to be used in the present study. An abundance of research has indicated that the SVT is a reliable and valid measure of listening and reading comprehension. The reliability and validity of the CAAS system is also supported by ample evidence. While reports of reliability evidence for phonological awareness tasks are scarce, the data that is available suggests that phonological awareness tasks are reliable and valid measures of phonological processing skill.

APPENDIX C

CALCULATION OF THE COMBINED ACCURACY/RESPONSE TIME INDEX

Sinatra and Royer (1995) have revealed a new procedure for combining accuracy and response time scores into a single index of performance. The procedure is as follows. An examinee's accuracy score is first converted to an inaccuracy score. In doing this, a high inaccuracy score and a high response time score both indicate poor performance, while a low inaccuracy score and low response time score both indicate better performance.

Next, an examinee's inaccuracy score is divided by the standard deviation obtained from the sample and his/her response time score is divided by its corresponding sample standard deviation. Each of these scores, which resulted from dividing by the standard deviation, is then squared. The scores are then added together and the square root is taken. The result is the combined index.

FOOTNOTES

1 The Counselor has an M.S. in Counseling with specialization in Special Education and has had 7 years of experience in the learning disability field.

2 It is acknowledged that the reliability of CAAS tasks as they currently exist would need to be investigated before the CAAS system could be used for diagnostic purposes.

3-4 One disabled subject was excluded from the analysis due to missing data on the phonological processing tasks.

5-6 One subject in the LD category was excluded from the analysis due to missing data on the phonological processing tasks.

7 Further classification of subjects into groups having only reading comprehension problems, only listening comprehension problems, or both problems was not possible since the documentation for most subjects often did not include information on both listening and reading comprehension.

8 One subject in the "reading problem" group was not included in the analysis due to missing data on the phonological processing tasks.

9 One subject in the "decoding and comprehension" group was not included in the analysis due to missing data on the phonological processing tasks.

BIBLIOGRAPHY

Aaron, P. G. (1987). Developmental dyslexia: Is it different from other forms of reading disability? Annals of Dyslexia, 37, 109-125.

Aaron, P. G. (1991). Can reading disabilities be diagnosed without using intelligence tests? Journal of Learning Disabilities, 24(3), 178-186, 191.

Aaron, P. G., Kuchta, S., & Grapenthin, C. T. (1988). Is there a thing called dyslexia? Annals of Dyslexia, 38, 33-49.

Adams, M. J., & Bruck, M. (1993). Word recognition: The interface of educational policies and scientific research. Reading and Writing, 5, 113-139.

Baddeley, A. D., Ellis, N. C., Miles, T. R., & Lewis, V. J. (1982). Developmental and acquired dyslexia: A comparison. Cognition, 11, 185-199.

Ben-Dror, I., Pollatsek, A., & Scarpati, S. E. (1991). Word identification in isolation and in context by college dyslexic students. Brain and Language, 40, 471-490.

Bradley, L., & Bryant, P. E. (1985). Rhyme and reason in reading and spelling. Ann Arbor: University of Michigan Press.

Bruck, M. (1988). The word recognition and spelling of dyslexic children. Reading Research Quarterly, 23, 51-69.

Bruck, M. (1990). Word-recognition skills of adults with childhood diagnoses of dyslexia. Developmental Psychology, 26, 439-454.

Bruck, M. (1992). The persistence of dyslexics' phonological awareness deficits. Developmental Psychology, 28(5), 874-886.

Bruck, M., & Treiman, R. (1990). Phonological awareness and spelling in normal children and dyslexics: The case of initial consonant clusters. Journal of Experimental Child Psychology, 50, 156-178.

Bryant, P. E., Maclean, M., & Bradley, L. (1990). Rhyme, language, and children's reading. Applied Psycholinguistics, 11, 237-252.

Bryant, P. E., Maclean, M., Bradley, L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection and learning to read. Developmental Psychology, 26(3), 429-438.

Carlisle, J. F., & Felbinger, L. (1991). Profiles of listening and reading comprehension. Journal of Educational Research, 84(6), 345-354.

Compton, D. L., & Carlisle, J. F. (1994). Speed of word recognition as a distinguishing characteristic of reading disabilities. Educational Psychology Review, 6(2), 115-140.

Cisero, C. A., & Royer, J. M. (1995). The development and cross-language transfer of phonological awareness. Contemporary Educational Psychology, 20, 275-303.

Cisero, C. A., Royer, J. M., Marchant, H., & Jackson, S. (1995, April). Can the Computer-based Academic Assessment System (CAAS) be used to identify reading disability?: A look at CAAS profiles of reading-disabled college students. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Cisero, C. A., Royer, J. M., Marchant, H., & Wint, F. T. (1994, April). Diagnostic identification of dyslexic and learning disabled college students using the Computer-based Academic Assessment System (CAAS). Poster presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. New York: Harcourt, Brace, Jovanovich.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334.

Dale, E., & O'Rourke, J. (1976). The living word vocabulary: The words we know. Field Enterprises Educational Corporation.

DiBenedetto, B., Richardson, E., & Kochnower, J. (1983). Vowel generation in normal and learning disabled readers. Journal of Educational Psychology, 75, 576-582.

Ellis, N., & Large, B. (1987). The development of reading: As you seek so shall you find. British Journal of Psychology, 78, 1-28.

Epps, S., Ysseldyke, J. E., & Algozzine, B. (1983). Impact of different definitions of learning disabilities on the number of students identified. Journal of Psychoeducational Assessment, 1, 341-352.

Evans, L. D. (1990). A conceptual overview of the regression discrepancy model for evaluating severe discrepancy between IQ and achievement scores. Journal of Learning Disabilities, 22, 406-412.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K., et al. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. Journal of Educational Psychology, 86(1), 6-23.

Fodor, J. (1983). Modularity of mind. Cambridge: MIT Press.

Forness, S. R., Sinclair, E., & Guthrie, D. (1983). Learning disability discrepancy formulas: Their use in actual practice. Learning Disability Quarterly, 6, 107-114.

Francis, W. N., & Kucera, H. (1982). Frequency analysis of English usage: Lexicon and grammar. Boston: Houghton-Mifflin.

Frankenberger, W., & Fronzaglio, K. (1991). A review of States' criteria and procedures for identifying children with learning disabilities. Journal of Learning Disabilities, 24(8), 495-500.

Gajar, A., Salvia, J., Gajria, M., & Salvia, S. (1989). A comparison of intelligence achievement discrepancies between learning disabled and non-learning disabled college students. Learning Disabilities Research, 4(2), 119-124.

Gough, P. B., & Tunmer, W. (1986). Decoding, reading, and reading disability. Remedial and Special Education, 7(1), 6-10.

Greene, B. A., Royer, J. M., & Anzalone, S. J. (1990). A new technique for measuring listening and reading literacy in developing countries. International Review of Education, 36, 57-68.

Hulme, C. (1988). The implausibility of low-level visual deficits as a cause of children's reading difficulties. Cognitive Neuropsychology, 5(3), 369-374.

Irlen, H. (1991). Reading by the colors: Overcoming dyslexia and other reading disabilities through the Irlen method. NY: Avery Publishing.

Kochnower, J., Richardson, E., & DiBenedetto, B. (1983). A comparison of the phonic decoding ability of normal and learning disabled children. Journal of Learning Disabilities, 16(6), 348-351.

Lewin, T. (1995, January 8). The learning disabled hit college. The New York Times Education Life. Section 4A, pp. 26-27, 29, 31, 52, 54.

Liberman, I. Y., & Liberman, A. M. (1990). Whole language versus code emphasis: A critical evaluation of the underlying assumptions and their implications for reading instruction. Annals of Dyslexia, 40, 51-76.

Lindgren, S. D., DeRenzi, E., & Richman, L. C. (1985). Cross-national comparisons of developmental dyslexia in Italy and the United States. Child Development, 56, 1404-1417.

Lovegrove, W. J., Martin, F., & Slaghuis, W. L. (1986). A theoretical and experimental case for a visual deficit in specific reading disability. Cognitive Neuropsychology, 3(2), 225-267.

Lundberg, I., Frost, J., & Peterson, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. Reading Research Quarterly, 23, 263-284.

Lynch, D. J. (1986). Reading comprehension performance as a function of individual differences in working memory. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Lynch, D. J. (1987). Reading comprehension under listening, silent and round-robin reading conditions as a function of text difficulty and working memory. Paper presented at the annual meeting of the American Educational Research Association, Washington, D.C.

Maclean, M., Bryant, P. E., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. Merrill-Palmer Quarterly, 33, 255-281.

Manis, F. R., Szeszulski, P. A., Holt, L. K., & Graves, K. (1988). A developmental perspective on dyslexic subtypes. Annals of Dyslexia, 38, 139-153.

Manis, F. R., Szeszulski, P. A., Holt, L. K., & Graves, K. (1990). Variation in component word recognition and spelling skills among dyslexic children and normal readers. In T. H. Carr & B. A. Levy (Eds.), Reading and its development: Component skills approaches (pp. 207-259). NY: Academic Press.

Manis, F. R., Szeszulski, P. A., Howell, M. J., & Horn, C. C. (1986). A comparison of analogy- and rule-based decoding strategies in normal and dyslexic children. Journal of Reading Behavior, 18(3), 203-218.

Morris, R. (1993). Issues in empirical versus clinical identification of learning disabilities. In G. R. Lyon, D. B. Gray, J. F. Kavanagh, & N. A. Krasnegor (Eds.), Better understanding learning disabilities (pp. 73-93, Baltimore, MD: Paul H. Brookes.

Morrison, F. J. (1984). Word decoding and rule-learning in normal and disabled readers. Remedial and Special Education, 5, 20-27.

Morrison, F. J. (1987). The nature of reading disability: Toward an integrative framework. In S. Ceci (Ed.), Handbook of cognitive, social, and neuropsychological aspects of learning disabilities (pp. 33-62). Hillsdale, NJ: Erlbaum.

Morrison, S. R., & Siegel, L. S. (1991). Learning disabilities: A critical review of definitional and assessment issues. In J. E. Obrzut & G. W. Hynd (Eds.), Neuropsychological foundations of learning disabilities (pp. 79-97). San Diego, CA: Academic Press.

Norušis, M. J. (1990). SPSS/PC+ Advanced Statistics Manual 4.0. The Netherlands: SPSS, Inc.

Olson, R. K. (1985). Disabled reading processes and cognitive profiles. In D. B. Gray & J. F. Kavanagh (Eds.), Biobehavioral measures of dyslexia (pp. 215-244). Parkton, MD: York Press.

Olson, R. K., Kliegl, R., Davidson, B., & Foltz, G. (1985). Individual and developmental differences in reading disability. In T. Waller (Ed.), Reading research: Advances in theory and practice (Vol. 4, pp. 1-64). London: Academic Press.

Pennington, B. F., Van Orden, G. C., Smith, S. D., Green, P. A., & Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexics. Child Development, 61, 1753-1778.

Perfetti, C. A. (1985). Reading ability. New York: Oxford University Press.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. Ehri, & R. Treiman (Eds.), Reading acquisition (pp. 1450174). Hillsdale, NJ: Erlbaum.

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. Just & P. Carpenter (Eds.), Cognitive processes in comprehension (pp. 141-183). Hillsdale, NJ: Erlbaum.

Perfetti, C. A., Beck, I., Bell, L. C., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. Merrill-Palmer Quarterly, 33(3), 283-319.

Posner, M. I., Boies, S. J., Eichelman, W. H., & Taylor, R. L. (1969). Retention of physical and name codes of single letters. Journal of Experimental Psychology Monographs, 79, 1-16.

Rasool, J. M., & Royer, J. M. (1986). Assessment of reading comprehension using the sentence verification technique: Evidence from narrative and descriptive texts. Journal of Educational Research, 79, 180-184.

Rayner, K., Murphy, L. A., Henderson, J. M., & Pollatsek, A. (1989). Selective attentional dyslexia. Cognitive Neuropsychology, 6(4), 357-378.

Reynolds, C. R. (1981). The fallacy of "two years below grade level for age" as a diagnostic criterion for reading disorders. Journal of School Psychology, 19(4), 350-358.

Reynolds, C. R. (1985). Measuring the aptitude-achievement discrepancy in learning disability diagnosis. Remedial and Special Education, 6(5), 37-55.

Royer, J. M. (1990). The Sentence Verification Technique: A new direction in the assessment of reading comprehension. In S. M. Legg and J. Algina (Eds.), Cognitive assessment of language and math outcomes (pp. 144-191). Norwood, NJ: Ablex.

Royer, J. M. (1995). The Sentence Verification Technique: A practical procedure for measuring listening and reading comprehension. Unpublished manuscript.

Royer, J. M., & Carlo, M. S. (1991). Assessing the language acquisition progress of Limited-English-Proficient Students: Problems and a new alternative. Applied Measurement in Education, 4, 85-113.

Royer, J. M., & Hambleton, R. (1983). Normative study of 50 reading comprehension passages that use the Sentence Verification Technique. Unpublished manuscript.

Royer, J. M., & Sinatra, G. M. (1994). A cognitive theoretical approach to reading diagnostics. Educational Psychology Review, 6(2), 81-113.

Royer, J. M., Carlo, M. S., Carlisle, J. F., & Furman, G. A. (1991). A new procedure for assessing progress in transitional bilingual education programs. Bilingual Review, 16, 3-14.

Royer, J. M., Carlo, M. S., & Cisero, C. A. (1992). School-based uses for the Sentence Verification Technique for measuring listening and reading comprehension. Psychological Test Bulletin, 5(1), 5-19.

Royer, J. M., Hastings, C. N., & Hook, C. (1979). A sentence verification technique for measuring reading comprehension. Journal of Reading Behavior, 11, 355-363.

Royer, J. M., Kulhavy, R. W., Lee, J. B., & Peterson, S. E. (1986). The Sentence Verification Technique as a measure of listening and reading comprehension. Educational and Psychological Research, 6, 299-314.

Royer, J. M., Lynch, D. J., Hambleton, R. K., & Bulgareli, C. (1984). Using the Sentence Verification Technique to assess the comprehension of technical text as a function of level of expertise. American Educational Research Journal, 21, 839-869.

Royer, J. M., Marchant, H., Sinatra, G. M., & Lovejoy, D. (1990). The prediction of college course performance from reading comprehension performance: Evidence for general and specific factors. American Educational Research Journal, 27, 158-179.

Royer, J. M., Sinatra, G. M., & Schumer, H. (1990). Patterns of individual differences in the development of listening and reading comprehension. Contemporary Educational Psychology, 15, 183-196.

Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. Journal of Child Psychology and Psychiatry, 16, 181-197.

Shepard, L. (1980). An evaluation of the regression discrepancy method for identifying children with learning disabilities. The Journal of Special Education, 14(1), 79-91.

Siegel, L. S. (1985). Psycholinguistic aspects of reading disabilities. In L. Siegel & F. Morrison (Eds.), Cognitive development in atypical children (pp. 45-65). New York: Springer-Verlag.

Siegel, L. S. (1988). Evidence that IQ scores are irrelevant to the definition and analysis of reading disability. Canadian Journal of Psychology, 42, 201-215.

Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. Journal of Learning Disabilities, 22(8), 469-478, 486.

Siegel, L. S. (1992). An evaluation of the discrepancy definition of dyslexia. Journal of Learning Disabilities, 25(10), 618-629.

Siegel, L. S., & Faux, D. (1989). Acquisition of certain grapheme-phoneme correspondences in normally achieving and disabled readers. Reading and Writing: An Interdisciplinary Journal, 1, 37-52.

Siegel, L. S., & Heaven, R. K. (1986). Categorization of learning disabilities. In S. J. Ceci (Ed.), Handbook of cognitive, social, and neuropsychological aspects of learning disabilities (Vol. 1, pp. 95-121). Hillsdale, NJ: Erlbaum.

Sinatra, G. M. (1989). Implementation and initial validation of a computer-based system for the assessment of reading competence. Unpublished Ph.D. dissertation, University of Massachusetts, Amherst, MA.

Sinatra, G. M., & Royer, J. M. (1993). The development of cognitive component processing skills that support skilled reading. Journal of Educational Psychology, 85, 509-519.

Sinatra, G. M., & Royer, J. M. (1995, April). Investigating a new method of combining response time and accuracy scores on measures of reading ability. Paper presented at the annual meeting of the Society for the Scientific Study of Reading, San Francisco, CA.

Slaghuis, W. L., & Lovegrove, W. J. (1985). Spatial-frequency-dependent visible persistence and specific reading disability. Brain and Cognition, 4, 219-240.

Snowling, M. (1980). The development of grapheme-phoneme correspondence in normal and dyslexic readers. Journal of Experimental Child Psychology, 29, 294-305.

Snowling, M. (1981). Phonemic deficits in developmental dyslexia. Psychological Research, 43, 219-234.

Spring, C., & French, L. (1990). Identifying children with specific reading disabilities from listening and reading discrepancy scores. Journal of Learning Disabilities, 23(1), 53-58.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. Reading Research Quarterly, 21(4), 360-406.

Stanovich, K. E. (1988a). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable difference model. Journal of Learning Disabilities, 21(10), 590-604, 612.

Stanovich, K. E. (1988b). The right and wrong places to look for the cognitive locus of reading disability. Annals of Dyslexia, 38, 154-177.

Stanovich, K. E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. Developmental Review, 10, 72-100.

Stanovich, K. E. (1991a). Conceptual and empirical problems with discrepancy definitions of reading disability. Learning Disability Quarterly, 14, 269-280.

Stanovich, K. E. (1991b). Discrepancy definitions of reading disability: Has intelligence led us astray? Reading Research Quarterly, 26(1), 7-29.

Stanovich, K. E. (1992). Speculations on the causes and consequences of individual differences in early reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), Reading Acquisition (pp. 307-342). Hillsdale, NJ: Erlbaum.

Stanovich, K. E. (1993). A model for studies of reading disability. Developmental Review, 13, 225-245.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. Journal of Educational Psychology, 86(1), 24-53.

Stanovich, K. E., Cunningham, A. E., & Cramer, B. B. (1984). Assessing phonological awareness in kindergarten children: Issues of task comparability. Journal of Experimental Child Psychology, 38, 175-190.

Stein, J. F., & Fowler, S. (1982). Ocular motor dyslexia. Dyslexia Review, 5, 25-28.

Venezky, R. L. (1970). The structure of English orthography. The Hague, Netherlands: Mouton.

Vogel, S. A. (1982). On developing LD college programs. Journal of Learning Disabilities, 15, 518-528.

Wint, F. T., Cisero, C. A., & Royer, J. M. (1995, April). Poor readers have different problems as revealed by the Computer-based Academic Assessment System (CAAS). Poster presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Woods, P. A., Sedlacek, W. E., & Boyer, S. P. (1990). Learning disability programs in large universities. NASPA Journal, 27(3), 248-256.

Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. Reading Research Quarterly, 23(2), 159-177.